HEALTH TECHNOLOGY ASSESSMENT

VOLUME 20 ISSUE 22 MARCH 2016
ISSN 1366-5278

The natural history of *Chlamydia trachomatis* infection in women: a multi-parameter evidence synthesis

Malcolm J Price, AE Ades, Kate Soldan, Nicky J Welton, John Macleod, Ian Simms, Daniela DeAngelis, Katherine ME Turner and Paddy J Horner

DOI 10.3310/hta20220

NHS
National Institute for Health Research

The natural history of *Chlamydia trachomatis* infection in women: a multi-parameter evidence synthesis

Malcolm J Price,[1] AE Ades,[2]* Kate Soldan,[3] Nicky J Welton,[2] John Macleod,[2] Ian Simms,[3] Daniela DeAngelis,[3,4] Katherine ME Turner[2] and Paddy J Horner[2,5]

[1]Institute of Applied Health Research, University of Birmingham, Birmingham, UK
[2]School of Social and Community Medicine, University of Bristol, Bristol, UK
[3]Public Health England (formerly Health Protection Agency), Colindale, London, UK
[4]Medical Research Council Biostatistics Unit, Cambridge, UK
[5]Bristol Sexual Health Centre, University Hospital Bristol NHS Foundation Trust, Bristol, UK

*Corresponding author

Declared competing interests of authors: KMET reports personal fees from Aquarius Population Health, outside the submitted work. PJH reports personal fees from Aquarius Population Health, grants, personal fees and non-financial support from Cepheid, personal fees from Crown Prosecution Service, personal fees from British Association for Sexual Health and HIV, grants from Mast Group Ltd, grants and personal fees from Hologic, outside the submitted work; in addition, PJH has a patent A sialidase spot test to diagnose bacterial vaginosis, issued to University of Bristol.

Declared competing interests of authors: none

Published March 2016
DOI: 10.3310/hta20220

This report should be referenced as follows:

Price MJ, Ades AE, Soldan K, Welton NJ, Macleod J, Simms I, *et al*. The natural history of *Chlamydia trachomatis* infection in women: a multi-parameter evidence synthesis. *Health Technol Assess* 2016;**20**(22).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (SciSearch®) and *Current Contents®*/ Clinical Medicine.

Health Technology Assessment

HTA MRC–NIHR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 5.027

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme or, commissioned/managed through the Methodology research programme (MRP), and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

This issue of the Health Technology Assessment journal contains a project funded by the MRC Infections and Immunity Board and promoted under the auspices of the Methodology Research Programme (an MRC NIHR Partnership), which is published in the journal by collaborative agreement due to the relevance of the technology being investigated.

The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the MRC, NETSCC, the HTA programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the MRC, NETSCC, the HTA programme or the Department of Health.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Abstract

The natural history of *Chlamydia trachomatis* infection in women: a multi-parameter evidence synthesis

Malcolm J Price,[1] AE Ades,[2]* Kate Soldan,[3] Nicky J Welton,[2] John Macleod,[2] Ian Simms,[3] Daniela DeAngelis,[3,4] Katherine ME Turner[2] and Paddy J Horner[2,5]

[1]Institute of Applied Health Research, University of Birmingham, Birmingham, UK
[2]School of Social and Community Medicine, University of Bristol, Bristol, UK
[3]Public Health England (formerly Health Protection Agency), Colindale, London, UK
[4]Medical Research Council Biostatistics Unit, Cambridge, UK
[5]Bristol Sexual Health Centre, University Hospital Bristol NHS Foundation Trust, Bristol, UK

*Corresponding author T.Ades@bristol.ac.uk

Background and objectives: The evidence base supporting the National Chlamydia Screening Programme, initiated in 2003, has been questioned repeatedly, with little consensus on modelling assumptions, parameter values or evidence sources to be used in cost-effectiveness analyses. The purpose of this project was to assemble all available evidence on the prevalence and incidence of *Chlamydia trachomatis* (CT) in the UK and its sequelae, pelvic inflammatory disease (PID), ectopic pregnancy (EP) and tubal factor infertility (TFI) to review the evidence base in its entirety, assess its consistency and, if possible, arrive at a coherent set of estimates consistent with all the evidence.

Methods: Evidence was identified using 'high-yield' strategies. Bayesian Multi-Parameter Evidence Synthesis models were constructed for separate subparts of the clinical and population epidemiology of CT. Where possible, different types of data sources were statistically combined to derive coherent estimates. Where evidence was inconsistent, evidence sources were re-interpreted and new estimates derived on a post-hoc basis.

Results: An internally coherent set of estimates was generated, consistent with a multifaceted evidence base, fertility surveys and routine UK statistics on PID and EP. Among the key findings were that the risk of PID (symptomatic or asymptomatic) following an untreated CT infection is 17.1% [95% credible interval (CrI) 6% to 29%] and the risk of salpingitis is 7.3% (95% CrI 2.2% to 14.0%). In women aged 16–24 years, screened at annual intervals, at best, 61% (95% CrI 55% to 67%) of CT-related PID and 22% (95% CrI 7% to 43%) of all PID could be directly prevented. For women aged 16–44 years, the proportions of PID, EP and TFI that are attributable to CT are estimated to be 20% (95% CrI 6% to 38%), 4.9% (95% CrI 1.2% to 12%) and 29% (95% CrI 9% to 56%), respectively. The prevalence of TFI in the UK in women at the end of their reproductive lives is 1.1%: this is consistent with all PID carrying a relatively high risk of reproductive damage, whether diagnosed or not. Every 1000 CT infections in women aged 16–44 years, on average, gives rise to approximately 171 episodes of PID and 73 of salpingitis, 2.0 EPs and 5.1 women with TFI at age 44 years.

Conclusions and research recommendations: The study establishes a set of interpretations of the major studies and study designs, under which a coherent set of estimates can be generated. CT is a significant cause of PID and TFI. CT screening is of benefit to the individual, but detection and treatment of incident infection may be more beneficial. Women with lower abdominal pain need better advice on when to seek early medical attention to avoid risk of reproductive damage. The study provides new insights into the reproductive risks of PID and the role of CT. Further research is required on the proportions of PID, EP and TFI attributable to CT to confirm predictions made in this report, and to improve the precision of key estimates. The cost-effectiveness of screening should be re-evaluated using the findings of this report.

Funding: The Medical Research Council grant G0801947.

Contents

List of tables

List of figures

List of abbreviations

ACCEPTt	Australian Chlamydia Control Effectiveness Pilot
BASHH	British Association for Sexual Health and HIV
BUGS	Bayesian inference Using Gibbs Sampling
BV	bacterial vaginosis
CDC	Centers for Disease Control
CEA	cost-effectiveness analysis
CI	confidence interval
ClaSS	Chlamydia Screening Studies project
CrI	credibility interval
CT	*Chlamydia trachomatis*
CT–	*Chlamydia trachomatis* negative, not *Chlamydia trachomatis* infected
CT+	*Chlamydia trachomatis* positive, *Chlamydia trachomatis* infected
D	posterior mean of the residual deviance
DAG	directed acyclic graph
DIC	deviance information criterion
ELISA	enzyme-linked immunosorbent assay
EP	ectopic pregnancy
FE	fixed effect
FOI	force of infection
FP	family planning
GP	general practitioner
GPRD	General Practice Research Database
GUM	genitourinary medicine
GUMCAD	Genitourinary Medicine Clinic Activity Dataset
HES	Hospital Episode Statistics
HFEA	Human Fertilisation and Embryology Authority
HIV	human immunodeficiency virus
HPA	Health Protection Agency
HR	hazard ratio
IUD	intrauterine device
IVF	in vitro fertilisation
KC-60	routine statistical returns on attendances at STI clinics in the UK
LCR	ligase chain reaction
LOR	log odds ratio
MCMC	Markov chain Monte Carlo
MIF	micro-immunofluorescence
MPES	Multi-parameter Evidence Synthesis
MRC	Medical Research Council
NAAT	nucleic acid amplification test
NCSP	National Chlamydia Screening Programme
NICE	National Institute for Health and Care Excellence
NSC	National Screening Committee
OR	odds ratio
p_D	The effective number of parameters, in a Bayesian hierarchical model
PEF	Population Excess Fraction
PID	pelvic inflammatory disease
POPI	Prevention of Pelvic Infection (trial)
RCT	randomised controlled trial
RE	random effect
RR	relative risk

RRR	relative risk reduction	WBDev	WinBUGS Development Interface
SD	standard deviation	WBDiff	WinBUGS Differential Interface
STD	sexually transmitted disease	WIF	whole cell immunofluorescence
STI	sexually transmitted infection	WinBUGS	Windows implementation of BUGS
TFI	tubal factor infertility		

Plain English summary

Chlamydia trachomatis (CT) infection is a cause of pelvic inflammatory disease (PID), ectopic pregnancy (EP) and infertility but is usually asymptomatic. The National Chlamydia Screening Programme (NCSP) offers sexually active young men and women annual screening for CT, but its effectiveness and value for money have not been definitely established. There has been no consensus on how much reproductive damage untreated CT infection can cause or on the best methods for estimating this.

The purpose of this project was to assemble all the evidence on CT infection, PID, EP and tubal factor infertility (TFI) to see if a consistent picture could be formed of the role of CT in causing reproductive damage. We looked at the worldwide literature on the risks of these outcomes following CT, and also at the UK evidence on CT prevalence, and routinely reported PID and EP seen in general practitioner surgeries, hospitals and sexually transmitted disease clinics, and at evidence on infertility from UK surveys.

We found a way of interpreting all this evidence which provided a single consistent set of estimates. It was confirmed that untreated CT infection posed a significant threat to reproductive health. Our findings show that screening is beneficial to the individual, but a focus on treating infections at the time they are acquired may have greater benefit.

The cost-effectiveness of the NCSP now needs to be reassessed, using these new estimates of how much reproductive damage CT infection can cause.

Scientific summary

Background and objectives

The National Chlamydia Screening Programme (NCSP) was initiated in 2003 and was operational throughout England by 2008. It offers screening for *Chlamydia trachomatis* (CT) infection annually, and on partner change, to sexually active men and women under 25 years attending general practitioner (GP) surgeries and other services. In 2012, over 1.7 million chlamydia tests were carried out in England among young people aged 15–24 years old. Assuming one test per person, this approximates to 35% of young women and 16% of young men being tested for chlamydia. The objectives of the programme (see www.chlamydiascreening.nhs.uk/ps/overview.asp) are to:

- prevent and control CT through early detection and treatment of infection
- reduce onward transmission to sexual partners
- prevent the consequences of untreated infection, principally pelvic inflammatory disease (PID), which can result in ectopic pregnancy (EP) and tubal factor infertility (TFI).

The effectiveness and cost-effectiveness of the programme has yet to be clearly established.

Treatment of current CT infection is considered to be both safe and effective, and screening has been proposed because the majority of incident CT remains asymptomatic. However, the cost-effectiveness and clinical effectiveness of screening depends primarily on the precise extent of reproductive morbidity caused by untreated CT, chiefly PID, EP and TFI, and on the proportion that can be prevented by screening.

The work reported here was funded by the Medical Research Council. The premise for the project was that there was no consensus on the quantitative risks of PID, EP and TFI following a CT infection, or on how to derive estimates from the evidence available, or even on what kind of evidence should be used. The objective, therefore, was to comprehensively assemble all the available evidence on the incidence and prevalence of CT in the UK, and the evidence from the various prospective and retrospective study designs from which quantitative relationships between CT, PID, EP and TFI can be derived, as well as routine sources of evidence on PID, EP and TFI in the UK, in order to assess the consistency of the different types of evidence, and, if possible, to provide a coherent, unified account of the clinical and population epidemiology of CT in the UK and its reproductive consequences in women.

Methods

Evidence sources were identified using 'high-yield' strategies, based on citations in recent cost-effectiveness analyses, reviews and research papers, and on the advice of the multidisciplinary group of investigators. New formal systematic reviews were not undertaken. Where routine UK data were used, this was from 2002, prior to the introduction of the NCSP.

In the interests of transparency and simplicity, the problem was broken down into separate but interlinked subproblems. These were: duration of asymptomatic CT; incidence, prevalence and duration of CT considered together; the risk of PID following CT infection; the incidence of PID and the proportion of PID attributable to CT; the cumulative incidence of PID, of repeat episodes and the prevalence of previous salpingitis; the relation between salpingitis and EP; the relation between salpingitis and TFI; and the relation between CT and TFI from serological case–control studies. Under each of these headings, multiple sources of evidence were assembled and their interpretation was reviewed. Models were estimated from the assembled data following a Multi-parameter Evidence Synthesis (MPES) approach, using Bayesian Markov chain Monte Carlo.

Where appropriate, we combined the evidence from multiple sources to form a single coherent set of estimates for multiple parameters (such as incidence, prevalence and duration). Wherever possible, we assessed the consistency of estimates derived from alternative evidence sources. Where evidence sources were in conflict, we attempted to identify alternative sets of assumptions, or alternative interpretations of the data sources, under which the different sources of data could be regarded as making consistent predictions for the model parameters. In such cases, where models and interpretations were based on post-hoc reasoning, this was highlighted, and any conclusions were considered as tentative and requiring further confirmation.

Results

The key results from each of the analyses are as follows.

Duration of asymptomatic Chlamydia trachomatis *infection in women (see* Chapter 4*)*

Evidence on CT duration in women was extraordinarily heterogeneous. However, the heterogeneity can be explained if studies of incident and prevalent infection are distinguished, and if one assumes that 'passive' infections clear over approximately 1 week. A model including such passive and 'real' infections gave an adequate fit. The evidence was also compatible with a three-rate model which includes fast clearing real infection as a result of a protective immune response in addition to passive infection and slow clearing real infection.

Incidence, prevalence, and duration of Chlamydia trachomatis *in the UK (see* Chapter 5*)*

It was shown that available evidence on CT incidence in the UK (infection rates and re-infection rates), appropriately calibrated to apply to the general population, was consistent with evidence on duration and prevalence. Key findings were:

- Approximately 77% [95% credible interval (CrI) 68% to 84%] of incident CT in women is asymptomatic.
- An asymptomatic CT infection has an average duration of 1.31 years (95% CrI 1.06 to 1.56 years).
- A CT infection (symptomatic and asymptomatic) has an average duration of 1.03 years (95% CrI 0.82 to 1.25 years).
- CT prevalence in women ranges from 8.4% per year in 16- to 17-year-olds to 0.8% in 30- to 44-year-olds. It was 5.2% (95% CrI 3.8% to 6.9%) in 16- to 24-year-olds and 2.1% (95% CrI 1.6% to 2.7%) in 16- to 44-year-olds.
- CT incidence ranges from 8.2 per 100 person-years in 16- to 18-year-olds to 0.8 in 30- to 44-year-olds. It was 5.0 per 100 person-years (95 CrI 3.5 to 7.1) in 16- to 24-year-olds and 2.1 per 100 person-years (95% CrI 1.5 to 2.8) in 16- to 44-year-olds.

Risk of pelvic inflammatory disease following Chlamydia trachomatis *infection (see* Chapter 6*)*

A Markov model was constructed which allowed for CT clearance as a 'competing risk' alongside the development of PID. Estimates of the proportion of incident CT that progresses to PID were generated, based on synthesised data from three trials, as well as estimates of the proportion of CT-related PID that could be prevented by screening on an annual basis. We explored the possibility that rates of progression to PID could be higher in the 3 months following infection, but the data available could not distinguish between one- and two-rate models:

- 14.9% (95% CrI 4.8% to 24.8%) of incident CT infections progress to symptomatic PID
- 17.1% (95% CrI 5.6% to 28.9%) of incident CT infections progress to PID (symptomatic and asymptomatic)
- 7.3% (95% CrI 2.3% to 14%) of incident CT infections progress to salpingitis
- In women aged 16–24 years who undergo screening at annual intervals, at best, 61% (95% CrI 55% to 67%) of CT-related PID and 22% (95% CrI 7% to 43%) of all-cause PID can be directly prevented.

Incidence of pelvic inflammatory disease and proportion of pelvic inflammatory disease attributable to chlamydia (see Chapter 7*)*

We established that the all-cause PID incidence as observed in the Prevention of Pelvic Infection (POPI) trial, taking account of the proportion that is asymptomatic (13%), is consistent with routinely collected data on PID from hospital, GP and genitourinary medicine clinic returns, taking account of the overlap between these data sets and the proportion of PID that is undiagnosed (64%).

- A pooled estimate of PID incidence in 16- to 24-year-olds, including diagnosed and undiagnosed PID, is 2.5 per 100 person-years (95% CrI 1.8 to 3.4), and in 16- to 44-year-olds is 1.8 per 100 person-years (95% CrI 1.3 to 2.5).
- 62.9% (95% CrI 57.8% to 67.4%) of PID episodes are in women aged > 24 years.

Several estimates of the proportion of PID due to CT, the Population Excess Fraction (PEF) (see *Chapter 3*), were compared:

- the PEF is four to six times higher in 16- to 19-year-olds than in 35- to 44-year-olds
- the preferred estimate of the PEF was 35.3% (95% CrI 10.5% to 68.5%) in 16- to 24-year-olds, and 19.7% (95% CrI 5.9% to 38.1%) in 16- to 44-year-olds, based on estimates of CT incidence, CT-related PID incidence and the CT-to-PID progression risk
- this is consistent with estimates derived from case–control studies, adjusted for under-ascertainment of CT infection, and with estimates of relative risk reduction from the POPI trial although uncertainty is high.

Cumulative incidence of Chlamydia trachomatis*-related and non-*Chlamydia trachomatis*-related pelvic inflammatory disease and salpingitis (see* Chapter 8*)*

Based on a Markov model of PID and salpingitis incidence and repeat episodes:

- A total of 42.9% (95% CrI 25.5% to 61.2%) of incident PID would be confirmed as salpingitis on laparoscopy.
- In women aged 35–44 years, 33.6% (95% CrI 25.4% to 43.1%) have experienced at least one episode of PID (all causes, diagnosed and undiagnosed), and 16.1% (95% CrI 9.0% to 24.7%) have experienced at least one episode of salpingitis (all causes, diagnosed and undiagnosed).

Comparison of ectopic pregnancy rates predicted from the Lund study with UK data on ectopic pregnancy incidence (see Chapter 9*)*

Although conception rates peak between the ages of 20 and 30 years, the proportion of conceptions that are EPs rises sharply with age, indicating its sensitivity to cumulative exposure to risk factors. The proportion of EP that is due to salpingitis was derived from two French case–control studies.

- 1.13% of all pregnancies in the UK are EPs
- an estimated 27% (95% CrI 11% to 46%) of EPs are due to salpingitis
- an estimated 4.9% (95% CrI 1.2% to 12.1%) of EPs are due to CT.

We derived predictions for UK EP conception rates based on the salpingitis-to-EP risks observed in the Lund study, and a parallel set of predictions were made for TFI. It was concluded, however, that the salpingitis-to-EP risks observed in the Lund study were too high to be consistent with UK data on EP, possibly because of changes in interuterine device use.

Comparison of tubal factor infertility rates predicted from the Lund study with UK infertility surveys (see Chapter 10*)*

- Prevalence of primary and secondary TFI in women aged 44 years was 1.08% (95% CrI 0.79% to 1.54%), based on UK infertility surveys.
- This is consistent with the salpingitis-to-TFI risks observed in the Lund study, if it is assumed that the TFI risks associated with *all* salpingitis, whether diagnosed or not, is the same as, or slightly lower than, the salpingitis-to-TFI risk observed in the Lund study.
- An estimated 29% (95% CrI 9% to 56%) of TFI is attributable to CT.

Proportion of tubal factor infertility attributable to Chlamydia trachomatis*, based on serological case–control studies (see* Chapter 11*)*

We developed a method for estimating the proportion of TFI cases due to CT from serological studies, with adjustment for the sensitivity and specificity of the serological assays. This was applied to a case–control study from the Netherlands. It was estimated that 45% (95% CrI 28% to 62%) of TFI was attributable to CT in this study, but this is likely to be an overestimate. There is a large body of evidence clearly demonstrating that CT is a significant cause of TFI.

Conclusions

The study has generated a set of estimates on chlamydia epidemiology, from its incidence, prevalence and duration of infection, all the way through to its role in PID, EP and TFI. These estimates are not only consistent with an extensive body of literature, and with fertility surveys and routine statistics on PID and EP, but they are also internally coherent. To achieve this, the study has produced a coherent set of interpretations of the key study designs.

Public health significance

- Our findings confirm that CT is an important cause of PID and TFI.
- The findings support the view that screening of prevalent cases prevents PID, but suggest that a greater emphasis should be placed on detection and treatment of incident CT infection, as part of an integrated programme for sexually transmitted infection treatment and control.
- Current guidance on PID management, regarding a low threshold for presumptive treatment with broad-spectrum antibiotics, should not be changed.
- Women with lower abdominal pain require advice on when to seek early medical attention to avoid risk of reproductive damage.
- Every 1000 CT infections in women aged 16–44 years, on average, gives rise to approximately 171 episodes of PID and 73 of salpingitis, 2.0 EPs and 5.1 women with TFI at age 44 years.

Limitations

The study has a limited scope: it has not covered dynamic models of CT transmission, and therefore cannot by itself fully inform cost-effectiveness analyses of screening. Neither costs nor impact on quality of life have been addressed. Chronic pelvic pain, CT infection in pregnancy, and the role of CT in neonatal pneumonia and conjunctivitis have not been covered.

Within its scope, the main limitations relate to the large number of assumptions that have been made, although these are assumptions that have been commonly made in the previous literature, in particular:

- the proportion of PID that is undiagnosed is the same, regardless of age and whether or not the PID is CT related
- the proportion of PID that would be confirmed as salpingitis on laparoscopy is the same, whether or not it is CT related
- the proportion of PID that would be confirmed as salpingitis on laparoscopy is the same, whether or not it is diagnosed
- the reproductive damage caused by salpingitis is the same whether it is CT related or not.

Recommendations for further research

Further research is recommended as follows:

- A suite of serological studies, based on routine health service activity, should be undertaken to estimate the causal role of CT in PID, EP and TFI, and how this might vary with age.
- Such studies would also offer the opportunity to gain up-to-date information on: PID referral patterns, leading to better estimates of PID incidence; the proportion of PID causing reproductive damage that is diagnosed, and the proportion that is silent; whether or not the CT-related PID is more or less likely to be associated with reproductive damage, and is more or less likely to diagnosed. Microbiological studies of the aetiology of reproductive damage following salpingitis could also be undertaken.
- Further dynamic modelling, within a MPES framework, so that *all* sources of evidence can be incorporated and checked for consistency, with appropriate uncertainty propagation. Further research may be required to develop methods of Bayesian computation capable of incorporating sexual network dynamics in disease transmission models.

Funding

Funding for this study was provided by the Medical Research Council grant G0801947.

Chapter 1 Background

Objectives

1. To set out the background to this research, and the context in which it was conceived.
2. To describe its motivation, objectives and the analytic strategy that has been adopted.
3. To set out the structure of the monograph.

Introduction

We begin with a brief description of the National Chlamydia Screening Programme (NCSP) in England, and the decision-making processes that led to it. Our account suggests that, at the time it began, *Chlamydia trachomatis* (CT) screening in the UK was not based on clear evidence of its effectiveness or cost-effectiveness. The present study, the questions it seeks to answer and the methodology it uses, originate in the early history of NCSP, and in the way in which evidence was used in the two UK studies[1,2] of the cost-effectiveness of screening for CT, published in 2007.

The National Chlamydia Screening Programme in England

The early history of the programme

The NCSP was set up in 2003 and was implemented in some way in every region of England by 2008. It is an opportunistic approach involving a combination of health and non-health screening venues, targeted at sexually active young people aged 15–24 years.[3] The objectives of the programme include both individual and population benefit.

1. Reduce CT incidence (population benefit). Detection and treatment of CT directly lowers CT prevalence and this will reduce the incidence of new cases in future.
2. Treat CT and prevent sequelae (individual benefit). Detection and early treatment of CT should prevent progression to sequelae.
3. Treat sexual partners of index cases.

The website (www.chlamydiascreening.nhs.uk/) provides further information on how the programme operates.

The precise evidence base for the NCSP has been controversial both before[4] and since its introduction.[5] In the UK, new screening programmes are normally examined by the National Screening Committee (NSC), based on a specifically commissioned systematic review and cost-effectiveness analysis (CEA) of candidate proposals for screening. The committee applies a set of criteria (see www.screening.nhs.uk/criteria) derived from the Wilson and Junger criteria for screening,[6] and considers the infrastructural implications of the screening and diagnostic tests, training in counselling and other services. In the case of CT, the decision to go ahead with a screening programme was originally based on the report from the Chief Medical Officer's Expert Advisory Group on *Chlamydia trachomatis.*[7] This group considered that the proposed programme met the necessary criteria, although others[5] have found otherwise. Although the report recommended screening in sexually active women aged < 25 years, women aged 25 years or over with a new partner in the previous 12 months, and high-risk men, there was a recognition that data on the costs, feasibility and acceptability of the programme were lacking, and these were to be addressed by pilot studies to be initiated the following year, in 1999. An early dynamic modelling study undertaken during this period by the Department of Health[8] may also have influenced the decision. It had concluded that screening could be cost neutral within 4 years.

In the event, following supportive experience during the pilot studies, CT screening formed an important plank in the English government's 2001 sexual health strategy[9] and has continued to do so ever since. The NSC reviewed the population-based screening programme in 2006 but did not recommend it.[10] The NSC may revisit this, or it may review opportunistic screening as currently practiced in England, in the future. Sheringham *et al.*[11] recently provided an insightful analysis of the development of policy leading to the NCSP.

Evaluating the benefits of Chlamydia trachomatis *screening*

As with any screening programme, evaluation of its effects can be undertaken in two ways: one method is experimental studies in which individuals are randomised to either screening and early treatment or to no screening, with outcomes assessed at appropriate later time points. At the time when the NCSP was set up, two randomised controlled trials (RCTs) had been conducted,[12,13] one of which, a large-scale study[13] on women considered to be at higher risk in the USA, appeared to show that the frequency of pelvic inflammatory disease (PID) could be approximately halved in the screened group. This trial[13] has been criticised by a number of commentators,[14–16] although the US Task Force for Prevention did not see fit to adjust the estimates of benefit[17] and we return to consider this more closely in later chapters. A second study, by Ostergaard *et al.*,[12] was relatively small. Both of these trials[12,13] studied an intervention and target population somewhat different from that proposed and implemented in England. Further RCTs have been published since the NCSP was introduced.

By 2009 there had been more reviews of the trial evidence than there had been trials.[18] Once again, although there was good evidence that screening could be effective in lowering the incidence of PID in high-risk populations, and also in high school students, some felt that there was a lack of evidence that screening would be successful in the general population.[18] Since then, the Prevention of Pelvic Infection (POPI) trial[19] has studied whether or not screening can lower the PID rate in women in higher education. The findings of this trial[19] were again positive. However, the results were considered to be less than definitive: the trial was somewhat underpowered because of difficulties with recruitment and lower-than-expected rates of PID.[20,21]

There have also been trials that have sought to estimate the effect of screening on population prevalence, rather than its effect on morbidity in the recruited population. Studies published to date have not delivered convincing results.[18] A stepped-wedge pilot in the Netherlands has reported no substantial fall in positivity among those screened over three screening rounds. The intended primary outcome of impact on population prevalence was aborted as a result of difficulties obtaining this measure.[22] A Danish trial[23] randomised women to a single postal invitation to be screened compared with no intervention, and followed up to PID, EP and infertility outcomes over a 9-year period. In view of the uptake of testing in both arms as part of routine care, and the low uptake of the offer of testing in the treatment arm, the result is not surprising.[24] In addition, the length of follow-up is likely to lead to considerable dilution of the relative risk (see *Chapter 3*, *Population Excess Fraction*). Results from the large-scale ACCEPt trial (Australian Chlamydia Control Effectiveness Pilot)[25] in Australia are still awaited.

A second approach to evaluation is by systematic review of the evidence and CEA. CEA applied to screening decisions can be informed directly by randomised studies, but is more often based on putting together the wide range of evidence on the population epidemiology of the condition, its clinical natural history with and without treatment, and – in the case of infectious disease programmes – models of the dynamic effects of interventions on future incidence. With sexually transmitted diseases (STDs), dynamic models must also make a range of assumptions about sexual partnership formation rates, duration of partnerships, their concurrency, and the impact of screening on partner recall and subsequent sexual behaviour, and so on. The advantage of evaluation by modelling, compared with RCTs, is that it is possible to consider the longer-term outcomes, beyond the trial, such as EP and tubal factor infertility (TFI). One can also study the expected dynamic effects of reducing prevalence. In 2007, two major CEAs of CT screening in the UK context were published.[1,2,15]

The results from the two CEAs were very different: one[2] concluded that screening was highly unlikely to be cost-effective (ClaSS, Chlamydia Screening Studies project), whereas the other concluded that screening was likely to be cost-effective, and most clearly so in men and women aged < 20 years if the risk of PID attributable to chlamydia was 10% or more. The fact that the two models produced different results is, in itself, not at all surprising. They were not examining exactly the same screening protocol, different structural assumptions were made about the natural history and different parameter values were assumed, particularly touching on transmissibility and sexual behaviour.[26,27]

However, what is surprising, particularly given the wide acceptance of systematic review as the methodology for analysing problems of this type, is that the disagreement between the two UK cost-effectiveness analyses[1,2] was not only about the parameter values, or even about how parameter values should be estimated from the available data, but also went as far as a conflict about what kinds of data should be used to inform parameters. Later in this chapter we illustrate the extent of the controversy concerning just one single parameter.

Population and individual benefits of screening

There are two distinct aspects to the epidemiological model underlying CEA of CT screening programmes: the dynamic effects on future incidence and prevalence of lowering prevalence by screening and treating a large proportion of the population, and the expected effect of screening and treatment of CT on progression to PID, EP and TFI. The two topics are not necessarily completely independent of each other, although it is important to establish that the two topics can be studied independently.

Although findings from the two UK screening models differed,[1,2] one clear finding that emerged from both CEAs was that the cost-effectiveness of CT screening was highly dependent on what was assumed to be the frequency of sequelae in untreated infection. The issue needs to be stated in two slightly different ways, each relating to the two objectives of screening:

1. If CT incidence is lowered by a given amount, how many PIDs, EPs and TFIs are prevented in the population as a whole?
2. In each case of CT detected by screening and treated, how much is the risk of PID, EP and TFI lowered?

The two questions are fundamentally different, and this has not always been fully appreciated. The first question invites us to consider how many further CT infections are prevented by treating each infected woman, recognising that each prevented CT infection can be associated with a number of averted sequelae. The second question relates to the benefits to the individual of detecting and successfully treating CT infection. This is the question that can be evaluated by randomising individuals in RCTs of screening versus no screening.

One way in which the questions of population benefit and individual benefit differ is that the former concerns *incident* infection, whereas the latter concerns *prevalent* infection. Women whose infection is picked up on screening are almost all asymptomatic (otherwise they would most likely have sought treatment already) and have already been infected for an unknown period. Trials of screening versus no screening show that treatment of asymptomatic *prevalent* infection does prevent subsequent PID. The issue of incident versus prevalent infections impacts on the interpretation of all the prospective studies of CT (see *Chapters 4* and *6*), as well as on the potential benefits of screening.

Difficulties with the evidence base for screening

In this section, we illustrate the extent of the controversy concerning what evidence should be used to inform models of CT screening and how it should be used.

Cost-effectiveness analysis models have agreed that the probability of PID following a CT infection is a key parameter in forming public health decisions on CT prevention and control. Although not all PID is caused by CT, the probability that an (untreated) episode of CT *causes* an episode of PID directly controls the downstream benefit of preventing a CT infection through screening or other preventative measures. Further, PID occupies a pivotal position among the various sequelae of infection in women because the consensus view[1,28–34] is that PID is a necessary precursor to a significant proportion of ectopic pregnancy (EP) and to virtually all TFI,[35] whether or not the PID is causally attributable to CT. Thus, increasing the probability that CT causes PID in any model should also lead to predictions of higher incidence of EP and TFI. (Although, this assumption was not built into the ClaSS study.[2]) The relationships between the clinical entities are shown in *Figure 1*.

Once again, given the central role of this parameter, it comes as a surprise that the two UK studies[1,2] sourced parameter estimates in completely different ways.

The ClaSS study[2] based their estimates on analyses of the Uppsala cohort,[36] a data linkage study involving all women aged 15–39 years, resident in Uppsala County, Sweden, in a 5-year period from 1985. Screening was offered by general practitioners (GPs), and laboratory screening records for the period were linked to hospital records of PID, EP and TFI. Setting aside the precise details of how estimates of progression rates were derived from the study, several difficulties can be enumerated, perhaps the most important of which is that women who were screened and found to be CT infected were treated. It is therefore not possible to draw conclusions about the progression rate to PID, EP and TFI in untreated women.

A very different approach was taken by the Health Protection Agency (HPA) study.[1] Here the investigators carried out scenario analyses around three estimates of the probability that an incident CT causes a PID: 1%, 10%, and 30%. The investigators proposed that the 10% PID rate should be considered the most plausible on the basis of an analysis of retrospective data. Their argument has two parts: first, overall PID incidence was between 1.5% and 2.4% per year, based on analyses of the General Practice Research Database (GPRD). Second, the proportion of total PID that could be attributed to chlamydia infection was approximately 30% according to studies that looked at evidence of current exposure in PID cases.[37] The HPA investigators then compared the supposed numbers of CT-related PIDs presenting to GPs with data on the number of incident CT cases. Whatever the validity of this approach, the use of *retrospective* data and the chain of reasoning via more than one type of study, contrasts with the use of a data linkage study in the ClaSS[2] model.

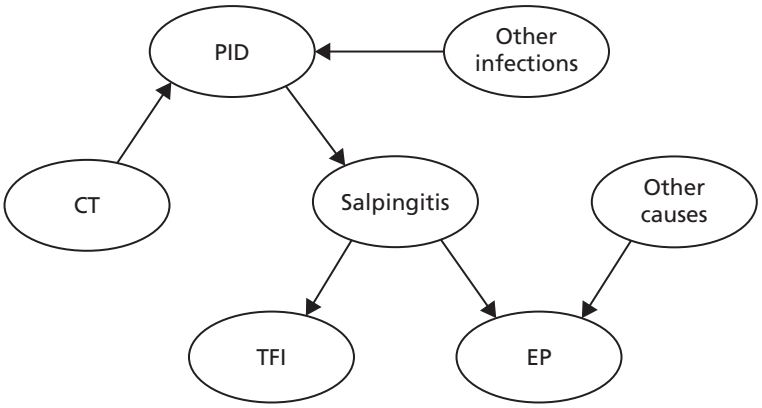

FIGURE 1 Relationships between CT infection and its major sequelae.

Subsequently, an authoritative series of papers on CT clinical and population epidemiology and the potential for CT screening appeared in the *Journal of Infectious Diseases* in 2010.[38] These were based on reports from an expert group set up by the US Centers for Disease Control and Prevention (CDC). Among them was a review of the research on the relation between CT and PID, co-authored by some of the authors of the ClaSS analysis. The review focused its attention not on linkage studies, *nor* on retrospective studies, but on a *third* type of study: prospective follow-up of untreated women with CT. Register linkage studies were recognised as possible sources of information, but were discounted for the reasons given above.

Even this does not exhaust the list of types of data that have been proposed. Another type of data was brought into play by van Valkengoed *et al.*[39] This was not prospective, retrospective or a linkage study, but simple routine statistics on the numbers of PIDs, EP and infertility being reported in routinely collected data. These estimates have been criticised,[40] and the reasoning behind them is questionable (see *Appendix 9*). However, although unconventional from a traditional Hierarchy of Evidence viewpoint,[41] the use of routine statistics to support inferences about causal effects is valid. Given an estimate of CT incidence in a population of known size, a register study of the number of PID episodes can put effective constraints on the CT-to-PID progression rate.

Summary of aims and objectives

Hopefully, our previous comment on the background to the NCSP, the way it was introduced and the difficulties that respected experts have had, not only in providing a CEA of screening, but even in identifying an appropriate methodology, explains the motivation behind this project. The aims of the proposed research, as they were stated at the outset, were to generate:

1. a consensus view of what existing data tell us about the natural history of untreated CT
2. a quantitative model of the natural history of CT, providing internally coherent and externally validated estimates of the prevalence and incidence of the complications of CT in the absence of treatment, suitable for modelling cost-effectiveness of interventions in the UK.

This monograph focuses attention on CT and its sequelae in women aged between 16 and 44 years of age, corresponding approximately to the period of reproductive activity. Many of the data sources relate to England, some to Great Britain and some to the UK. However, our general presumption is that, in the absence of any evidence to the contrary, there is no reason to believe that we would reach different conclusions for different nations within the UK, and conclusions are therefore intended to be relevant to the UK as a whole.

Any estimates of incidence and prevalence, both of CT and of its sequelae, must refer to a specified time, as well as a specified place. We have targeted the year 2002, which is immediately before the beginning of the NCSP. Where possible, routine data have been from that year, and, although it has not been possible to source all evidence from that year, we have interpreted evidence as being relevant to that year, unless otherwise stated.

We now explain the analytic strategy we adopted to achieve these aims.

Analytic strategy

In a previous section we have just picked out a single parameter, the risk of PID following CT, to show first that previous investigators have not simply inferred different parameter values from the same set of available data, but that they have sourced their parameter estimates from completely different kinds of data. This is by no means the only example that could be cited. These fundamental divergences of approach between previous investigators lie behind the analytic strategy we have adopted, which is essentially to look at *all the* sources of evidence, to see if they are consistent with each other, and to produce a set of internally consistent and coherent estimates, based on all the data sources – or at least all of those that we believe provide unbiased estimates of their target parameters. This is a large and complex undertaking, not only because of the range of types of study that could bear on each parameter, but also because the parameters are all related. For example, if we were to entertain the belief that CT incidence was at the high end of its credible interval (CrI), we should be committed to also believing that CT prevalence is higher and/or duration of infection is lower; also that there is a higher incidence of CT-related PIDs, a higher incidence of CT-related EP episodes and a higher prevalence of CT-related TFI. Therefore, when we examine the evidence about any one of these parameters we should examine the evidence about all of them, because we can learn about each one by looking at any of the others. The relationships between the key clinical entities are shown in *Figure 1*.

To meet this challenge we have adopted the following analytic strategy:

1. work through the natural history of CT and its sequelae in a logical sequence, starting with CT incidence, prevalence and duration, and ending with TFI
2. at each stage, review the different kinds of evidence available, prospective, retrospective and routine data, and review previous approaches to deriving parameter values from these evidence types
3. formulate a model for the evidence at each stage, and synthesise the evidence under that model, examining the consistency of estimates with the data
4. examine the consistency of the entire set of estimates overall.

Statistical combination of evidence

To carry out this programme, we have relied on Multi-parameter Evidence Synthesis (MPES). This is an approach to model building and evidence synthesis[42,43] that has properties that will prove valuable in the present context. It recognises that epidemiological studies do not necessarily generate direct evidence on the parameters of interest, but instead can inform *functions* of parameters. It has the technical means to combine information on parameters and functions of parameters within a single coherent structure; it offers standard statistical approaches to checking model fit, and specifically for determining whether or not different sources of evidence are in conflict regarding the information they contribute to any specified parameter; all data sources are formally incorporated via their likelihood, and parameter uncertainty, sampling uncertainty, and variation owing to heterogeneity is readily propagated through the evidence network.

Multi-parameter Evidence Synthesis is an essentially Bayesian methodology and it has been implemented using the popular Bayesian Markov chain Monte Carlo (MCMC) package WinBUGS version 1.4.3 (Medical Research Council Biostatistics Unit, Cambridge, UK).[44] But we emphasise that the use of Bayesian MCMC is not a necessary feature of our approach to evidence synthesis. Bayesian MCMC in WinBUGS makes it relatively easy for investigators to fit complex models, and the adoption of simulation-based Bayesian computation also greatly facilitates uncertainty propagation. However, similar results could be achieved, in principle, though perhaps with some difficulty, using Frequentist methods.

The idea of statistical combination of information on different functions of parameters appears to have been independently rediscovered a number of times in the last 80 years, with applications in physics, environmental risk control, deterministic dynamic models of population change, and in a range of computational frameworks, in both Frequentist and Bayesian forms.[45–48] The term ‘multi-parameter

synthesis' was coined by Victor Hassleblad,[49] one of the originators of the Confidence Profile Method,[50] a programme of research in clinical and descriptive epidemiology that was the immediate precursor to MPES and which made particular use of bias adjustment.

Besides review papers,[42,43,51] there are a number of applications of MPES to infectious disease epidemiology, often oriented to screening problems.[52–56] A relatively simple application, and perhaps the best introduction, setting out all of the main properties of MPES, is a stylised application to prenatal human immunodeficiency virus (HIV) testing.[57]

Multi-parameter Evidence Synthesis can be viewed as a purely technical extension of the single parameter meta-analysis, as described in numerous textbooks and tutorial papers,[58–60] to multi-parameter problems. But it is probably more accurate to see it as a philosophically different approach to evidence and evidence synthesis.

Coherence of estimates

An important property of this approach to synthesis is its internal *coherence*. The generation of epidemiological evidence tends to be fragmented: separate studies are made of CT prevalence, CT incidence and CT duration, and separate estimates are published. But, in truth, these three quantities are known to be related: *prevalence* = *incidence* × *duration*. Estimates of any one of these parameters need to be consistent with estimates of the other two *and* with whatever data inform any of them. In the same way, studies of CT incidence, progression rates to PID, EP and TFI should produce results that are mathematically related to the results from retrospective studies of the role of CT in PID, EP and TFI, and to routine statistics. MPES is simply a methodology in which these relationships can be explored in a formal way, while taking account of the full extent of uncertainty due to sampling variation, additional sources of variation and any concerns that might come under the heading of 'data quality'.

There is, in fact, nothing very new or original about putting different pieces of information together to see if they 'add up'. Many chlamydiologists have carried out 'back of the envelope' calculations with exactly this objective in mind, to see whether information on incidence and prevalence, put together with information on progression to PID, EP or TFI could be consistent with routine data or not. Or asking whether routine data, put together with information on the proportion of PID, EP or TFI attributable to CT could be consistent with CT incidence and prevalence information or not. Indeed, many calculations of this type have been published.[5,39,40,61,62] The programme of work described in this monograph uses MPES to examine exactly the same ideas, but within the framework of a formal statistical method, and perhaps in a more complete and systematic way.

Role of interpretation

Note, however, that MPES is not simply a 'synthesis exercise'. The synthesis is carried out under a very specific model of how the different data sources were generated. There is, therefore, as in any scientific exercise, an element of interpretation, and therefore of judgement. This arises in several ways. First, goodness-of-fit tests may not be able to distinguish between alternative models, forcing investigators to use their judgement. But a far more serious element of subjectivity lies in the fact that our overall judgement about the mechanisms that generate a set of data may depend greatly on our interpretation of just one or two studies. It is notoriously difficult, for example, to know if confounding variables or effect modifiers are present, or how great their impact is. Although conflicts between evidence sources may alert us to the presence of confounders and effect modifiers, the use of MPES does not remove the problems they create. We have therefore been explicit about where there is a subjective element to model choice, and particularly where there is an element of data-driven (post hoc) re-interpretation of evidence. In these instances, we follow the usual practice of describing the results as 'hypothesis generating', and we prioritise research recommendations in this light.

Overview of this monograph

In *Chapter 2* we present a brief account of the natural history and epidemiology of CT and its sequelae. The objective is to explain the relationships between the clinical entities, and the definitions we have adopted. *Chapter 3* sets out the methodology used throughout the monograph. This covers methods for literature identification, and gives some further details of the MPES approach, such as sensitivity to priors, technical aspects relating to the use of WinBUGS software, and our methods for assessing model fit and model selection.

The rest of the monograph takes the successive aspects of the natural history of CT and PID in turn and delivers a 'mini-synthesis' of specified subsets of the data. This approach allows us to gather together and synthesise all the direct evidence on each parameter, and then to build it incrementally into the overall model. We adopt an incremental strategy, in preference to putting every piece of evidence straight into a single overarching model. It may seem at first sight that this defeats the very advantages of MPES. However, the advantage of the incremental approach is that it commits us to a view on each parameter that can be examined and debated in reference to the existing literature. If we had combined all the data from the outset, it would be much more difficult to understand conflicts between the results and previous work.

Figure 2 provides a simple schematic guide to the contents of *Chapters 4–12*, in terms of the relationships between the clinical entities (see *Figure 1*). We begin at the start of the natural history process with a synthesis of data on the duration of CT in asymptomatic women (see *Chapter 4*). This is followed (see *Chapter 5*) by a synthesis of information of prevalence and incidence in the UK with the duration data.

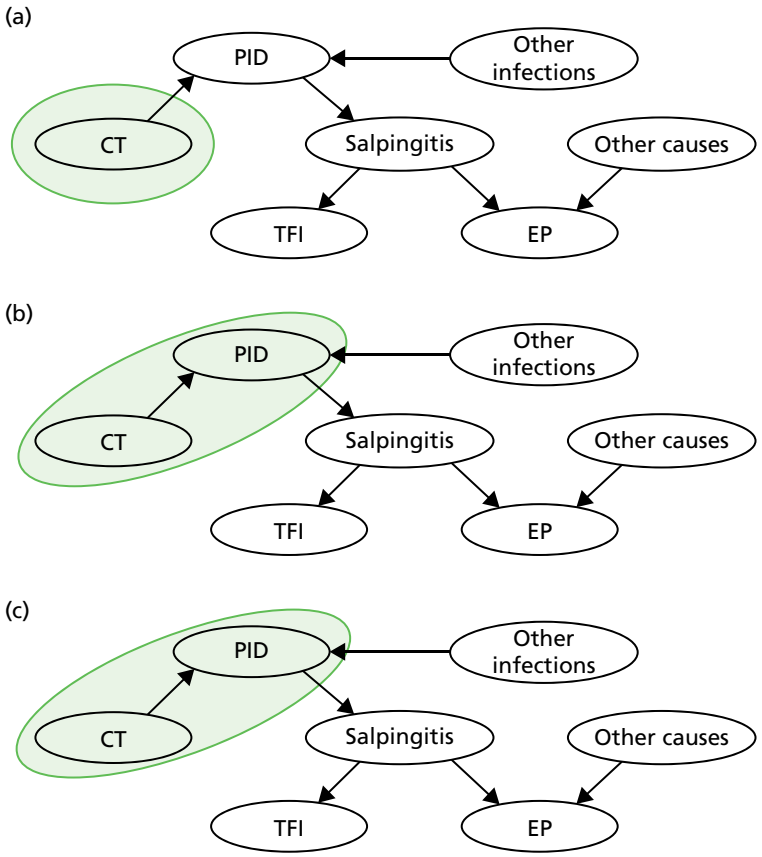

FIGURE 2 Guide to *Chapters 4–12* of the report. (a) *Chapters 4* and *5*, incidence, prevalence and duration; (b) *Chapter 6*, CT to PID progression; (c) *Chapter 7*, CT and PID; (d) *Chapter 8*, cumulative PID and salpingitis; (e) *Chapter 9*, salpingitis and EP; (f) *Chapter 10*, salpingitis and TFI; (g) *Chapter 11*, TFI and CT; and (h) *Chapter 12*, summary. (*continued*)

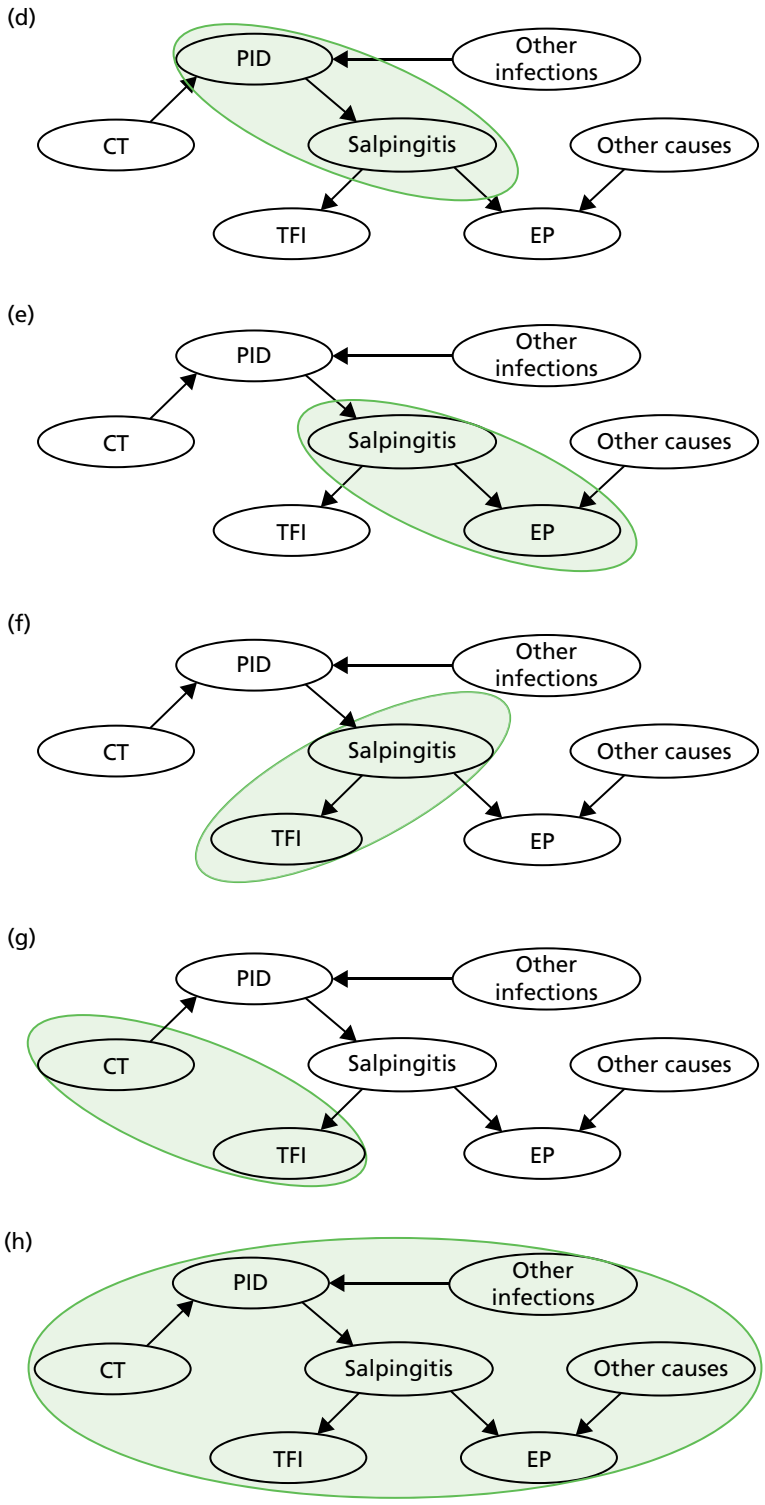

FIGURE 2 Guide to *Chapters 4–12* of the report. (a) *Chapters 4* and *5*, incidence, prevalence and duration; (b) *Chapter 6*, CT to PID progression; (c) *Chapter 7*, CT and PID; (d) *Chapter 8*, cumulative PID and salpingitis; (e) *Chapter 9*, salpingitis and EP; (f) *Chapter 10*, salpingitis and TFI; (g) *Chapter 11*, TFI and CT; and (h) *Chapter 12*, summary.

This allows us to establish the consistency of incidence, prevalence and duration data, and to go forward with a coherent analysis of these parameters. The next chapter (see *Chapter 6*) examines and re-analyses information from prospective studies, particularly randomised studies, following women with untreated CT forward to observe the occurrence of PID. This is used to generate estimates of the probability a CT episode causes PID. In *Chapter 7*, we examine the relationship between CT and PID from a retrospective viewpoint: how much PID is caused by CT. This requires us also to consider the incidence of PID.

There is strong evidence that the women with PID, who have an elevated risk of EP and TFI, are those in whom salpingitis can be confirmed by laparoscopy. In these women, EP and TFI risk increases with the number of PID episodes. *Chapter 8* presents a Markov model in which information on PID incidence and CT re-infection rates is used to generate predictions on the proportion of women who have experienced salpingitis followed by further episodes of PID. In *Chapter 9* we use these together with data from prospective and retrospective studies of salpingitis and PID to compare the predicted occurrence of EP to the rates of EP observed in the UK. *Chapter 10* takes the same form as *Chapter 9*, but deals with TFI rather than EP. *Chapter 11* derives sensitivity and specificity estimates for some new CT serology assays to adjust reported CT seroprevalence in a study comparing TFI cases and controls. We then derive estimates of the proportion of TFI attributable to CT. Finally, in *Chapter 12* we review the previous chapters, identifying where inconsistencies between estimates exist, and suggesting how these might be resolved.

Chapters 8–12 should be considered to be somewhat more tentative than *Chapters 4–7*. There has been little previous work looking at these different evidence sources, and the aetiological mechanisms relating CT and PID to EP and TFI are less well studied. As a result, there is a larger element of post-hoc reasoning as we attempt to identify what assumptions need to be made about the evidence sources in order to make them consistent, rather than the preferred more objective approach of assembling evidence and then checking consistency under a well-understood model of how the evidence sources are generated. *Chapter 12* ends with a tentative set of final estimates, an outline of study limitations, and research priorities.

Summary

1. At the time the NCSP was introduced, the evidence base for the cost-effectiveness of CT screening against its aims of reducing infection rates and reducing sequelae was mixed and questionable.
2. Two major UK studies[1,2] of the cost-effectiveness of CT screening sourced critical parameter estimates from completely different types of study, indicating a lack of consensus about what kinds of evidence should be used to inform natural history parameters.
3. Our analytic strategy is to look at *all* sources of evidence: retrospective, prospective and routine data.
4. MPES is used to assess the consistency of different types of evidence, and to develop an internally coherent model of the natural history and population epidemiology of CT.
5. The use of MPES does not remove the subjective elements in model choice and evidence interpretation.

Chapter 2 Natural history and epidemiology of *Chlamydia trachomatis*

Objectives

To:

- provide an introductory account of CT, and its natural history, clinical and population epidemiology in the UK in women
- explain what is known about the relations between CT, PID, salpingitis, EP and TFI, explaining features of their clinical and descriptive epidemiology in the UK
- provide definitions of clinically diagnosed PID and salpingitis, and to explain the concepts of undiagnosed symptomatic PID, and asymptomatic or silent PID.

Chlamydia trachomatis infection

Chlamydia trachomatis is an intracellular bacterial pathogen and a major cause of genital and eye disease. It comprises three human biovars. One causes trachoma, the most common form of blindness worldwide, and particularly prevalent in Africa. A second is the cause of lymphogranuloma venereum, a sexually transmitted infection (STI), mainly of gay men, in the developed world. The third biovar, consisting of serovars D to K, is the topic of this monograph. It is a major cause of PID, EP and TFI in women, epididymitis in men, and neonatal conjunctivitis and neonatal pneumonia in newborns.[3,28,63,64]

Chlamydia trachomatis is transmitted by oral, vaginal or anal sex, and can also be transmitted from mother to newborn during a vaginal delivery.[3,63,64] Diagnosis is generally based on nucleic acid amplification tests (NAATs) carried out on cervical or vaginal swabs in women, urethral swabs in men or on urine samples.[3,63,64] NAAT has largely replaced culture or enzyme immunoassays for diagnostic testing.[3,63,64] However, culture played a key role in earlier epidemiological studies, and the lower sensitivity of culture, 60–80%,[64,65] needs to be taken into account when interpreting the earlier research results.

The infection can be treated effectively by antibiotics, such as azithromycin or doxycycline, but it is often asymptomatic, especially in women.[3,63,64] Re-infection by the same partner is an important issue in prevention: partner notification rates can be improved and risks of re-infection reduced by providing index cases with prescriptions intended for their partners, or with the medication itself.[66–71]

Control and prevention

There appears to be little consensus on approaches to control and prevention. In the USA screening is recommended for non-pregnant women aged < 25 years, and for older women with risk factors (e.g. those who have a new sex partner or multiple sex partners).[63] The technical report from the European Centre for Disease Control 2008 lists the chlamydia-control activities in Europe, ranging from none to an organised CT screening programme.[16] In Sweden, a change in infectious disease laws in 1988 required doctors to offer testing and treatment to those with suspected chlamydia. This policy continues in five other countries. It should probably be considered as opportunistic testing for selected asymptomatic individuals, rather than as a formal screening programme.[5,16] Only two countries are detailed as having an organised CT screening programme: England and the Netherlands.[16] The opportunistic screening approach adopted in England is described in *Chapter 1*. In the Netherlands, a pilot register-based screening

programme for people aged 16–29 years was recently undertaken (2007–10).[72] Participation rates were low, < 20%, and decreased over three rounds of register-based screening.[22]

Natural history and epidemiology

Chlamydia trachomatis infection

Approximately 75% of incident infections in women are asymptomatic.[73] Although studies of the duration of asymptomatic infection have produced a wide range of estimates – between 1 and 18 months (see *Chapter 4*) – it is evident that the infection can clear spontaneously in the absence of treatment as a consequence of the adaptive and innate immune responses.[38,74]

The risk of transmission of infection from one episode of sexual intercourse is estimated to be between 10% and 20%.[75,76] Given the high sensitivity of current NAATs for detecting CT,[8–10] and the fact that non-host DNA and sperm can be recovered from the female genital tract up to 7 days following sexual intercourse, this strongly implies that at least some women who test CT positive (CT+) are probably only passively infected.[77,78] This is consistent with the observations of Joyner *et al.*[79] and Geisler *et al.*[80] that a proportion of women testing CT+ without treatment become CT NAAT-negative within 2 weeks.

Prevalence of CT can be studied by population surveys, using NAATs on genitourinary specimens. As with most STIs, incidence and prevalence are highest among young people, and decline steeply with age.[81] Infection is more common among the more sexually active individuals, with the greatest number of sexual partners.[8–10,81]

Pelvic inflammatory disease

Definition of pelvic inflammatory disease

Pelvic inflammatory disease comprises a spectrum of upper genital tract inflammatory disorders among women, which includes any combination of endometritis, salpingitis, tubo-ovarian abscess and pelvic peritonitis.[82] The principal clinical and economic significance of CT lies in the fact that it is a causal agent of PID and the further damage to reproductive health that follows from salpingitis.[33,83,84] Confusion can arise with the term PID. Historically, it was often used interchangeably with the term 'acute salpingitis' as detected at laparoscopy, which, historically, was the gold standard for diagnosing PID.[85,86] More recently endometritis, which is associated with salpingitis but can be detected by endometrial sampling, has been equated with PID because of the unacceptability of performing laparoscopy on large numbers of women with minimal symptoms.[85–87]

Pelvic inflammatory disease is not easy to diagnose and the criteria for a clinical diagnosis of PID have changed over time.[33] In the original Lund studies[33] the minimum criteria for diagnosis were: (1) lower abdominal or pelvic pain of < 3 weeks; (2) a purulent vaginal discharge diagnosed on microscopy or painful intermenstrual bleeding; and (3) increased motion tenderness of the uterus and adnexa on bimanual examination. In the UK national PID guideline 2011,[83] recent onset of lower abdominal pain in association with local tenderness on bimanual examination is now considered sufficient to establish a diagnosis and initiate treatment. This is because it is now recognised that many women with salpingitis have subtle or mild symptoms.[63,83] Even when present, the clinical symptoms and signs of PID lack sensitivity and specificity for detecting women with salpingitis.[83] Because of the difficulty of diagnosis and the potential for damage to the reproductive health of women (even by apparently mild PID) and the benefit of early antimicrobial therapy, health-care providers are now advised to maintain a low threshold for the diagnosis of PID.[83]

Laparoscopic diagnosis of PID, which identifies women with salpingitis, is no longer undertaken in women with a clinical diagnosis of PID. However, clinical information can be used to classify PID as 'possible',

'probable' and 'definite' PID, based on Hager's criteria.[88,89] *Table 1* details the clinical definitions adopted by Taylor-Robinson *et al.*,[88] based on a modified version of the Hager criteria, and which were used to estimate the proportion of women with salpingitis following a clinical diagnosis of PID. This classification is often used in clinical trials such as POPI, and in studies of patient data such as the GPRD database, although in recent work using GPRD 'probable' PID includes chronic disease as well.[19,88,90] As the probability that patients diagnosed with PID based on clinical presentation have salpingitis decreases, the likelihood of other diagnoses, such as endometriosis, irritable bowel syndrome and functional pain, increases.

For the purposes of this report we define clinically diagnosed PID as women with 'probable' and 'definite' diagnostic criteria. Women with 'possible' diagnostic criteria may be treated for PID depending on to where they present, which is probably less likely in primary care compared with departments of genitourinary medicine (GUM) as there is greater expertise in managing PID in the latter setting. Diagnostic thresholds have changed over time as the role of PID in reproductive damage has been recognised. Accordingly, the proportion of women with PID regarded as having 'probable/definite' or 'possible' PID diagnostic criteria is likely to have increased over time, with fewer women remaining symptomatic and undiagnosed.[19,90] This probably began in the mid to late 1990s in the UK.

Infections associated with pelvic inflammatory disease

Pelvic inflammatory disease is not exclusively caused by CT, with about 30–40% of PID cases being CT+[37,82] (see *Chapter 7*). A range of STIs have been implicated, particularly gonorrhoea and, more recently, *Mycoplasma genitalium*.[63,83,91] It is also likely that PID can be caused by micro-organisms associated with bacterial vaginosis (BV), which is commonly present in women with PID.[63,83,92] These micro-organisms include *Gardnerella vaginalis* and anaerobes (including *Prevotella*, *Atopobium* and *Leptotrichia*),[83,93] which are probably sexually transmissible but would not be classed as STIs.[92,93] Anaerobes are isolated more often from women with severe PID. As a consequence, broad-spectrum antibiotic therapy is recommended for PID to cover *Neisseria gonorrhoeae*, CT, and a variety of aerobic and anaerobic bacteria commonly isolated from the upper genital tract in women with PID.[83]

Bacterial vaginosis is a common condition affecting women of reproductive age.[94,95] It is a condition characterised by vaginal flora imbalance, in which the normal plentiful lactobacillus is scarce and other anaerobic bacteria abundant.[94–96] It is present in between 12% and 25% of women.[94,95] Women are usually asymptomatic and if symptoms are present these consist of vaginal discharge with or without an odour.[63] Only some BV-associated bacteria have been associated with PID, and it remains to be established whether or not these bacteria are indeed causally implicated in reproductive damage.[63]

TABLE 1 Definition of clinical diagnosis of PID

PID diagnosis	Definition
Definite	Recent-onset lower abdominal pain; cervical excitation with or without uterine/tenderness, with adnexal tenderness with or without mass, and clinically unwell, and LGTI
Probable	Recent-onset lower abdominal pain with (1) cervical excitation/uterine tenderness and adnexal tenderness or (2) cervical excitation/uterine tenderness or adnexal tenderness with evidence, LGTI or clinically unwell
Possible	Recent-onset lower abdominal pain and only one of the following: (1) cervical excitation/uterine tenderness or (2) adnexal tenderness or Chronic lower abdominal pain, not clinically unwell and no evidence of LGTI and only one of the following: (1) cervical excitation/uterine tenderness or (2) adnexal tenderness

LGTI, lower genital tract infection.
Definitions based on Taylor-Robinson *et al.*[88]

Undiagnosed pelvic inflammatory disease and asymptomatic pelvic inflammatory disease

Over 50% of PID episodes in the UK are treated in primary care, but it is also treated in STI clinics and in hospital. Data from the GPRD, Hospital Episode Statistics (HES), and KC-60 returns from STI clinics for 2002 are shown in *Table 2*, which is based on the 'definite/probable PID' definition.[90] We have assumed that women admitted for inpatient treatment of PID in the HES data, and women diagnosed at departments of GUM by clinicians trained in diagnosing PID, are 'probable' or 'definite' cases. An unknown proportion of diagnosed PID episodes meeting these definitions will be recorded under more than one of these locations, although the study by Nicholson *et al.*[97] using the GPRD database suggests that approximately one-third of cases of probable or definite PID diagnosed in primary care have been treated elsewhere.

These figures on diagnosed PID incidence certainly underestimate the true incidence of PID episodes. It is widely accepted that a relatively high proportion of PID is undiagnosed. This is inferred from historical studies of TFI. PID is considered to be a necessary precursor of TFI,[98] and yet a high proportion of women with TFI have no history of clinical PID.[99–102] Wolner-Hanssen[103] found that only 34% of TFI cases reported a previous diagnosis of PID. However, only 11% reported *never* having had clinical symptoms.[103] The implication is that a large proportion of the PID that causes TFI is undiagnosed, but it is relatively unusual for PID which causes reproductive damage to be completely asymptomatic. We therefore distinguish *diagnosed PID*, by which we mean PID meeting the 'probable/definite' criterion, *undiagnosed symptomatic PID* and *asymptomatic or silent PID* (see *Chapter 7*).

Therefore, even if we maintain a single diagnostic criteria, such as 'probable/definite', across all studies, and these definitions were consistently applied – which we know is not the case – it would still be the case that the proportions of PID that would be ascertained would depend on the study design. For example, the routine PID statistics in *Table 2* depend on patient self-referral, whereas in a prospective study of women attending a STI clinic, or in the context of a screening trial, a higher proportion of PID meeting the same definition may come to light (see *Chapter 7*).

Current guidance on pelvic inflammatory disease management

Recognising the difficulty of achieving a definitive diagnosis, the British Association for Sexual Health and HIV (BASHH) guidelines[83] emphasise a low threshold for empiric treatment, to lower the risk of subsequent EP and infertility. The same advice is seen in primary care guidelines and the National Institute for Health and Care Excellence (NICE) Clinical Knowledge Summaries.[104,105] However, the latter cites multiple partners, a recent new partner, young age and previous history of PID or a STI as risk factors for PID. The NICE Clinical Knowledge Summary[104] explicitly states that 'PID is almost always due to a sexually transmitted disease'. This contrasts somewhat with BASHH guidelines,[83] which acknowledge that only 25% of cases of can be accounted for by CT or *N. gonorrhoeae.* US guidelines highlight young age as a risk factor but avoid citing any behavioural risk factors.[63] All guidance advocates early use of broad-spectrum antibiotics.

TABLE 2 Number of incident cases of PID in England, 2002

Age (years)	HES	GPRD[a]	GUM[b]	Female population
16–19	1233	5083	3212	1,199,600
20–24	3101	8842	4399	1,519,100
25–34	9756	14,932	3919	3,502,100
35–44	10,526	9609	1388	3,795,600

GPRD, General Practice Research Database; GUM, genitourinary medicine; HES, Hospital Episode Statistics.
a Definite and probable PID as defined in French *et al.*[90]
b Data by age not available for 2002, so we assume that the age distribution for these data was the same in 2002 as in 2009.

Age profiles of Chlamydia trachomatis *and pelvic inflammatory disease*

In this section we briefly compare the age profiles of CT and PID, to draw attention to an important, but generally overlooked, issue in the natural history of CT. Estimates of CT prevalence are approximately fourfold higher in women aged 16–24 years than women aged 25–44 years.[106] The age profile of incident PID is very different. HES show an increasing trend in PID incidence rates with age (see *Table 2*). Evidence from the GPRD shows incidence rates that are higher in the age group 20–24 years, but the difference in PID incidence rates across ages is far less than the difference in chlamydia prevalence. The age profile in PID presenting in STI clinics is more similar to the age profile of incident CT, in that there is a relatively high proportion of young women. However, of the three sources, these make up the smallest proportion of diagnosed cases.

There are three plausible contributory reasons for these discrepancies in age profile. First, the PID diagnosis rates may increase with age. There is some evidence that the severity of PID might increase with the number of previous PID episodes. The proportion of women who become infertile because of tubal scarring from adhesions, and the proportion of pregnancies that are ectopic, appears to increases multiplicatively with number of previous PID episodes,[33] which is likely to be positively correlated with age, and it is likely that more severe PID has a higher chance of being diagnosed and hospitalised. This hypothesis is supported by the dramatic increase with age of hospital-diagnosed PID.

A second possibility is that the probability that a case of CT causes an episode of PID may be higher in older women, perhaps owing to more previous episodes of CT, other STIs and PID.[107]

Finally, the proportion of PID that is caused by CT may decline with age (see *Chapter 7*), possibly because of differences in the age distributions of the other risk factors for PID. Facultative and anaerobic bacteria are frequently isolated from women with PID; although often present in the normal vaginal flora, they occur more often and in increased concentrations, in women with BV,[82] which has been associated with PID. Interestingly, isolation of endogenous bacteria from the upper genital tract appears more common in older women and may be associated with severe suppurative disease, including tubo-ovarian abscess and recurrent PID.[82,108] Although BV does not appear to be associated with age, it is associated with increasing lifetime number of sexual partners.[94,109]

These three different explanations would have quite different public health implications. Analyses in subsequent chapters throw some light on the issues, but we do not believe it can be resolved on the basis of existing evidence, and this is addressed in the research recommendations in *Chapter 12*.

Salpingitis

Much of our knowledge of the impact of PID on reproductive health is based on the Swedish Lund study,[33,34,84,110] which recruited PID cases admitted to hospital between 1960 and 1984, and followed them forward to observe reproductive outcomes over a median 8-year period. The study was based on laparoscopic examination. Women with macroscopic inflammation of the fallopian tubes, 'salpingitis', were distinguished from PID cases with no macroscopic inflammation at laparoscopy. The more severe the macroscopic salpingitis, the greater the risk of adverse reproductive outcomes. Remarkably, the incidence of TFI and EP in the control group, those admitted to hospital with PID but with no salpingitis on laparoscopy, appeared to be no different from the general population, even though a lower genital tract infection was identified in the majority of cases.[33] It is likely that many of these women would have had endometritis or microscopic inflammation of the fallopian tubes but some will have had other pathology or causes for their pelvic pain.[83,111,112] Incidence of EP and TFI among women with salpingitis depended on age, the number of clinical PID episodes – and thus presumably on the number of salpingitis episodes – and on the severity of the salpingitis on admission (see *Chapters 9* and *10*).

These results raise the question: what proportion of clinically diagnosed PID is accompanied by an underlying macroscopic salpingitis? Unfortunately for our purposes, laparoscopy is no longer a routine procedure in examination of PID, so it is not possible to match up the definitions of PID and salpingitis

employed in the Lund study[33,34,84,110] to current practice or to recent research studies. However, we can deduce from the fact that they were hospitalised that the cases of PID recruited into the Lund study[33,34,84,110] would be at the more severe end of the spectrum. What is known is that the proportion of clinical PID cases with salpingitis fell during the course of the Lund study[33,34,84,110] from an initial 80% down towards 60% among those recruited at the end of the study.[110] In the analyses we present in *Chapters 10* and *11*, we have assumed that the proportion of PID cases with salpingitis has continued to fall.

Finally, there is little quantitative information on whether the organism causing PID impacts on the probability or extent of reproductive damage. In the Lund study[33,34,84,110] it was observed that the severity of CT salpingitis at laparoscopy was generally greater than would be expected from the relatively benign clinical picture.[101] For PID with salpingitis (including that caused by CT), delay of treatment for more than 3 days from onset of symptoms increases the likelihood of tubal damage.[84] Therefore, if the clinical symptoms of CT-related PID appear to be less severe than gonococcal PID, it is possible that women with chlamydia PID may present later than women with gonococcal PID,[101,113–115] and as a result suffer more reproductive damage. The same comments apply to women with asymptomatic PID.

Ectopic pregnancy

In a normal pregnancy the fertilised egg implants in the uterus, where it divides and develops into an embryo and, eventually, a fetus. An ectopic pregnancy is one in which the egg implants outside the uterus, most often in the Fallopian tubes. EP is an important cause of maternal morbidity and, occasionally, mortality.[116] It usually presents with acute symptoms including pelvic pain and vaginal bleeding, although increasingly EPs are diagnosed before the onset of symptoms, allowing early, conservative treatment. Prevention of maternal morbidity and mortality relies on early diagnosis and appropriate management, which is almost invariably in a specialist setting.[116]

Among the common causes of EP are fallopian tube damage because of salpingitis, surgery and smoking.[116] In a normal pregnancy, cilia in the Fallopian tubes carry the fertilised egg towards the uterus. If the cilia are damaged, the egg may become lodged in the tube. PID is an inflammatory process that can result in the development of scar tissue in the tubes. If both tubes become entirely blocked then fertilisation cannot take place at all. This constitutes a single common pathway leading from salpingitis to both EP and to TFI (which will be exploited in the modelling in *Chapters 9* and *10*).[98,116]

Data on overall EP rates in England are based on HES (HES code O00).[117] If these are compared with conception rates (maternities plus abortions),[118] it is evident that the EP rate, defined as the number of EPs divided by the number of conceptions, has remained remarkably constant between 2000 and 2009 (*Figure 3*), at around 1% of conceptions. Similar rates have been reported in France, although there was

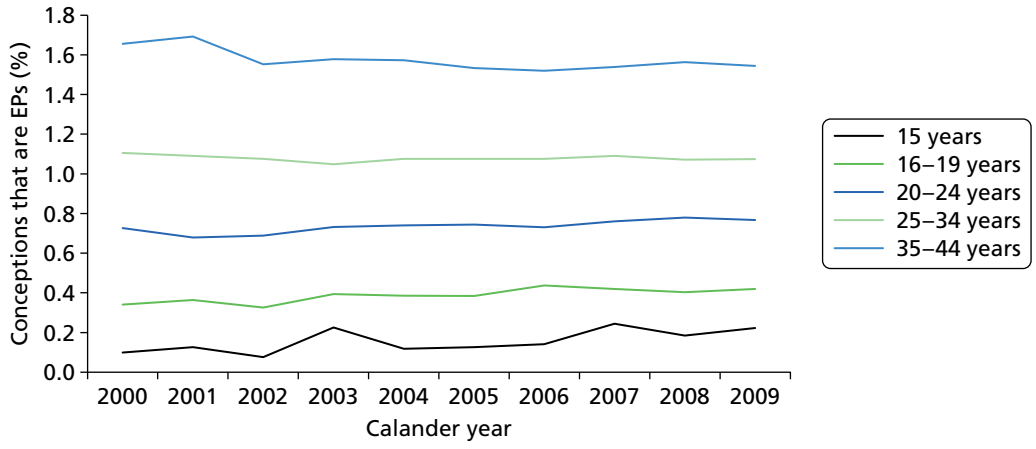

FIGURE 3 Proportion of conceptions that result in EP, by year and age group.

evidence that a fall in intrauterine device (IUD)-related EPs had masked a rise in EPs caused by salpingitis.[119] EP rates reported in the USA have been somewhat lower, but, again, have remained stable over recent years.[120]

Although the overall conception rate falls steeply from the age of 30 years onwards, the proportion of conceptions that are EPs rises with age (see *Figure 3*). The same phenomenon has been observed many times.[120–122] This points to the role of cumulative exposure to the causal factors, whether these are smoking or tubal damage, which can only increase with age.

The fact that EP risk links to *cumulative* exposure to risk factors, including salpingitis, rather than to, say, incident exposure, is critical to understanding the role of PID.[123] Egger *et al.*[121] report an ecological analysis of EP rates and CT prevalence by age and location in Uppsala county, Sweden. They establish that the correlation between CT prevalence and EP rates is strongest in 20- to 24-year-olds, and that it becomes increasingly weak in the 25- to 29-year-old, 30- to 34-year-old, and finally the 35- to 39-year-old groups. At the same time the regression of EP rate against CT prevalence becomes increasingly steep, which would not be expected if variation in EP rates was driven by variation in CT incidence. However, these phenomena can be readily explained if cumulative salpingitis incidence is the variable determining EP rates. Most CT occurs in the younger age groups, but current incidence and prevalence become increasingly poor predictors of cumulative exposure to CT as age increases. On the other hand, the regression against recent incidence becomes increasingly steep with age – although with increasing error about the regression – because cumulative incidence is higher per unit of current prevalence. These ecological observations provide further support for the conclusions we draw from *Figure 3* in the previous paragraph.

Another important feature of the relation between salpingitis and EP is the strong 'dose–response' relationship reported in prospective studies following women with salpingitis. The Lund studies,[33,110,124] in particular, followed women for an average of approximately 8 years from hospitalisation for PID. The proportion of women whose first subsequent pregnancy was an EP increased systematically with severity as determined by laparoscopy, and also with the number of previous salpingitis episodes.[124] Among the women who subsequently conceived, those with two, or three or more previous episodes of salpingitis had a non-linear increase in risk of EP (see *Chapter 9*). A similar pattern of findings was reported for TFI.[33] A record linkage study in the USA following high-risk women with a diagnosed CT infection over an average 6-year period again reported a similar increase in risk of EP, but this time in relation to the number of previous CT infections rather than the number of PID episodes.[107] These studies constitute some of the strongest evidence for a causal link between salpingitis (PID) and EP. Of particular significance is the fact that the dose–response relation is exactly what would be expected on the basis of *cumulative* exposure to risk factors. In addition, the similarity of the 'dose–response' effect attaching to CT and salpingitis suggests that the risk of EP may be approximately the same whatever the cause of the salpingitis, although the agreement could be coincidental if the aetiology changes with age in a way that approximately cancels out any differences.

Tubal factor infertility

Women are usually considered 'infertile' if they have tried to have a baby for ≥ 1 year and have been unsuccessful. Defined in this way, infertility is clearly not an absolute condition, as only 85% of couples succeed in conceiving within 1 year, whereas 95% may succeed within 3 or 4 years.[125] This immediately reveals a major difficulty in the study of infertility, as apparently minor changes in definition readily generate threefold changes in the number considered infertile.

Infertility is a complex construct: the infertile population includes both couples who are sterile, and couples who are subfertile. The former cannot conceive, but the latter can, if given enough time. Population surveys of infertility in the UK[126–128] distinguish between primary unresolved failure to conceive, between 2% and 3.5%, representing women who have not conceived at the end of their reproductive life, and secondary infertility, which includes women who have conceived at least once, but who subsequently

experience difficulty in conceiving. Secondary infertility must be included in our study because it is quite possible for salpingitis to be acquired by women who have already had children.

Surveys reporting secondary infertility have reported the numbers who have experienced at least two years of being unable to conceive.[126–128] According to one study,[126] the secondary infertility group included women who had already had one or two, and in one case eight, previous children. The 2–3.5% estimates of secondary infertility from these surveys are therefore probably an overestimate, but the extent of the overestimate depends on the distribution of duration of infertility.

A further problem is the difficulty of estimating the proportion of women who are infertile at any particular age. Not only is it necessary to ensure that their infertility has lasted sufficiently long to meet a specified definition, but infertility can only manifest itself, and thus be diagnosed, in women who are trying to become pregnant, although the underlying condition may have been present for a long time. This contrasts with the study of EP where reliable age-specific incidence data are available from routine hospital statistics.

To avoid these problems in this monograph we have followed a strategy that has been adopted in a number of epidemiological studies of infertility in the UK,[126–129] which is to focus attention on the proportion of the population who are infertile, which specifically means unable to conceive, at the end of their reproductive life, which we take to be the age of 44 years (see *Chapter 10*).

Tubal factor infertility is one of several causes of female infertility (*Figure 4*).[130] According to a 2006 survey by the Human Fertilisation and Embryology Authority (HFEA), 22% of those being treated by in vitro fertilisation (IVF) have TFI (*Table 3*), but a further 30% of infertility in women undergoing treatment is unexplained.[131]

The Lund studies[124] established the key relationship between PID and TFI, which parallels the relationship between EP and TFI. A higher rate of TFI was observed only in PID cases with confirmed salpingitis on laparoscopy, and the risk increased with the number of subsequent PID episodes. Critically, none of the 601 control group women with PID who were not confirmed as having salpingitis went on to have TFI. A difference between EP and TFI, however, is that TFI can occur only following an infective process,[35] namely salpingitis; EP on the other hand has a multifactorial aetiology.

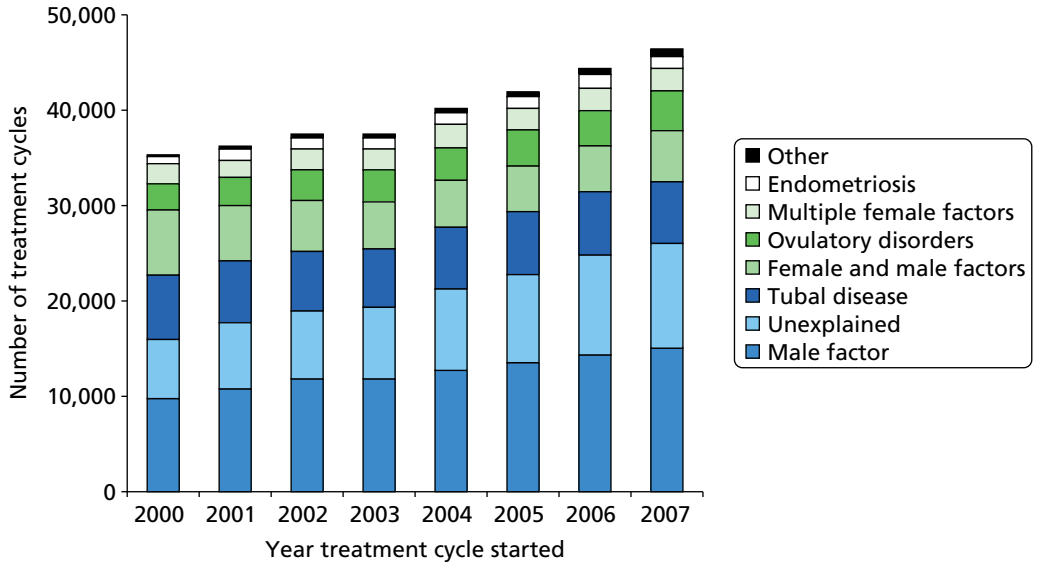

FIGURE 4 Types of infertility treated 2000–7. Source HFEA.

TABLE 3 Tubal factor infertility by age, 2002

Age (years)	Number of cycles in women with TFI (B)	Average number of cycles per woman with TFI (C)	Number of women with TFI adjusted for mean number of cycles: (D) = (B)/(C)	Number of cycles in women with other cause infertility (E)	Average number of cycles per woman with other-cause infertility (F)	Number of women with other-cause infertility adjusted for mean number of cycles: (G) = (E)/(F)	Proportion of women treated in HFEA with TFI: (H) = (D)/((D) + (G))
18–34	4550	2.20	2071	16773	2.43	6916	0.23
35–37	2247	2.51	896	8618	2.80	3076	0.23
38–39	1192	2.70	442	4404	2.92	1508	0.23
40–42	858	2.81	305	3663	3.13	1171	0.21
43–44	216	2.97	73	1102	3.07	359	0.17
45–50	86	3.37	26	767	3.17	242	0.10

Source: HFEA.

There is an extensive literature on the role of CT in TFI. Much of it consists of retrospective studies demonstrating a higher prevalence of serum antibody to CT in TFI cases compared with non-TFI controls. Because of the poor sensitivity and specificity of the assays, the imprecise characterisation of their diagnostic accuracy, and the variation in cut-offs, these studies have not generated reliable quantitative estimates of the role of CT in TFI. However, a number of these retrospective studies have reported marked differences in the titre distributions in antibody-positive TFI cases and controls (see *Chapter 11*). The higher prevalence of high titres in TFI cases constitutes powerful evidence of the causal role of CT.

Chronic pelvic pain

Chronic pelvic pain is a well-recognised complication of PID, which has a significant morbidity.[33,132] It is defined as non-menstrual pain of ≥ 6 months' duration that is severe enough to cause functional disability or require medical or surgical treatment. Chronic pelvic pain is common in women, with an estimated prevalence of 3.8% in women aged 15–73 years in the USA. Often the aetiology is not clear. There are many physical disorders of the reproductive tract, gastrointestinal system, urological, musculoskeletal system and psychoneurological system that may be associated with chronic pelvic pain in women,[132] and these may often co-exist. Although adhesions resulting from pelvic inflammatory disease are diagnosed in about 25% of women undergoing investigation, it is not clear whether these are always the cause of the pain.[132] The prevalence of adhesions in women with chronic pelvic pain will vary depending on the population studied. Most information is obtained from women attending secondary and tertiary care, who are likely to experience more severe forms and may not be representative of women with chronic pelvic pain in the community. An accurate diagnosis of pelvic adhesions is only possible following laparoscopy and, even then, it is not possible to ascertain the cause, as the infectious agent will have usually resolved prior to investigation.

We have therefore chosen not to investigate chronic pelvic pain because of the paucity of data on the role of PID, and the lack of evidence on the contribution of CT.

Premature delivery, neonatal pneumonia and conjunctivitis

There is conflicting evidence that CT may result in miscarriage and/or premature labour. One hypothesis is that adverse pregnancy outcomes may only occur following primary infection. No studies have investigated this since 1990. In view of the limited and conflicting data on adverse pregnancy outcomes, we decided not to include this in our work plan.

Chlamydia trachomatis is transmitted to approximately half of children of mothers infected with CT and about 10–50% of these develop neonatal conjunctivitis and/or pneumonia. This was recently reviewed by Thorne[133] for the NSC. Data in the UK are based on those presenting to hospital and thus biased towards severe and prolonged cases. Neonatal infection with chlamydia is usually asymptomatic, and the most common manifestations (conjunctivitis and pneumonia) are non-specific. Furthermore, chlamydial disease is usually of a mild to moderate nature and easily treated or is often self-limiting. For these reasons, perinatal and neonatal outcomes have not been included in this study.

Summary of clinical definitions and assumptions

1. The monograph focuses on:

    i. the mean duration, incidence, and prevalence of CT
    ii. the causal relationship between CT and PID
    iii. the relation between PID and salpingitis, by which we mean macroscopic inflammation of the fallopian tubes
    iv. the relationships between salpingitis and EP, between salpingitis and TFI, and between TFI and CT.

2. Premature delivery, neonatal outcomes and chronic pelvic pain are not covered in this report.
3. When referring to clinically diagnosed PID, we are referring to 'probable' and 'definite' PID as previously defined.[88,90]
4. We distinguish between diagnosed PID, undiagnosed symptomatic PID, and asymptomatic or silent PID.[103]

Chapter 3 Methods

Objectives

To set out methods used for:

1. evidence identification
2. Population Excess Fractions (PEFs)
3. evidence synthesis
4. model critique and model choice.

Introduction

This chapter sets out the methods used in evidence identification, evidence synthesis modelling and model critique. We provide only generic descriptions of our approach to these components of the project. The actual statistical synthesis models themselves, and the specific forms that the model critique took, are set out in detail in each of the subsequent chapters. We also include a section defining PEF and outlining the different ways PEFs can be estimated, because these feature in several chapters.

Evidence identification

The task of identifying data sources to inform the full range of CT natural history parameters from CT incidence through to EP incidence and TFI prevalence more closely resembles the task faced by investigators populating a decision-analytic model in a limited period of time than the task of systematic review. Although systematic review methods are considered ideal, it is uncommon for them to applied comprehensively in a multiple-parameter setting.[134] We have therefore relied on 'high-yield' and 'snowballing' strategies[135] that efficiently identify a large number of relevant data sources from a small number of cited references rather than on exhaustive systematic review. It is recognised that this introduces a risk of missing relevant sources. However, there is limited evidence comparing the sensitivity of the two approaches, particularly in the context of natural history parameters.[135] Nor is it clear that 'sensitivity' is an appropriate criterion, as the issue is not whether absolutely every literature source has been identified, but whether further searching would reveal anything that would change the conclusions.[135,136]

The search strategy was therefore to identify relevant literature cited in a series of dynamic modelling and cost-effectiveness analyses,[1,2,20,74,98,106,137–145] and review papers.[20,61,62,74,98,106,139,142–147] Cost-effectiveness analyses were used because their authors are obliged to have parameters with specified values, and must therefore cite the sources they used. A second tier of identification was based on a citation search on selected epidemiology papers identified in the first round. In addition the collaborative group undertaking the project included epidemiologists and STD clinicians with a very wide range of research experience in most aspects of chlamydia epidemiology, treatment and referral pathways in the UK, and routine and ad-hoc surveys on chlamydia and its sequelae in the UK. This formed an 'in-house' expert group who were consulted on the completeness of the literature identification, as well as on the interpretation of the data. Finally, in our study of the duration of CT infection (see *Chapter 4*) we consulted a panel of internationally recognised experts to ask if they knew of further papers that we had missed.

An important series of review papers published in *Journal of Infectious Disease* 2010[38] provided a particularly valuable resource. These were from an expert group set up by the CDC, Atlanta, Georgia, which included many of the world's acknowledged leaders in various fields of chlamydia research. The CDC expert group employed a more formal systematic review methodology. We have not made a formal study of whether or not these reviews identified papers that had not been identified by our searches, and we therefore cannot claim that the CDC project validates our approach. However, we believe that their work minimises the risk that we have overlooked information that is capable of having a material effect on the analyses we have produced.

The MPES approach, explained below, allows us to combine many different types of information in order to draw conclusions. For example, as we shall see in subsequent chapters, studies of different types may form a 'chain of evidence',[50] which can be put together to inform a *prospective* relation between CT and TFI. The results can then be compared with *retrospective* evidence on the relation between CT and TFI, which can also be derived by combining several kinds of data source. The question of which is most likely to provide an accurate estimate must be viewed with an open mind, bearing in mind both the relevance of the studies and their internal validity.

This contrasts somewhat with standard systematic review methodology, which tends to favour prospective over retrospective studies, and is not obviously compatible with reasoning via chains of evidence. For example, a systematic review of the causal effect of CT on TFI trawled through several thousand abstracts and identified a single study, which was then deemed to be inadequate, with the conclusion that no quantitative estimates could be derived.[145] This conclusion seems unwarranted as it would effectively rule out both retrospective evidence *and* estimates derived from a chain of evidence. In addition, systematic review would not allow certain kinds of evidence which are essential to our project, such as routine statistics, to be considered at all.

As noted above, a risk of missing sources of evidence would still remain even if we attempted to devise a formal systematic review on each parameter in every model. The strength of the MPES methodology, described below, is that it can combine direct evidence on model parameters with independent evidence on functions of model parameters. This raises further questions on whether or not evidence sources have been missed, as we know of no formal methods to define all the functions of parameters on which data *might* exist, let alone search strategies that could efficiently identify literature on each function.

However, it will become clear in subsequent chapters that, whatever the advantages and disadvantages of our approach to literature identification, we have combined and cross-checked many times more information on chlamydia, its incidence, prevalence and sequelae than previous researchers. Although we cannot be certain that every piece of relevant information has been incorporated, we believe that it is unlikely that anything has been missed that could have a material impact on our conclusions. It is in any case easy to incorporate further data within the framework we have created.

Details of included studies and reasons for study exclusion are set out separately for each chapter or in the appendices. Where appropriate, appendices are also used to explain how the data used in our analyses were derived from the published information. Further appendices set out the actual data as it was used in the data analyses, as well as the program code used to analyse it.

Population Excess Fractions

Population Excess Fractions feature in several places in the monograph. We outline the definition and estimation methods, as this has been an area of confusion in epidemiological and lay literature.[148] The PEF is the proportional reduction in disease risk that would be achieved by eliminating the exposure of interest from the population, assuming the exposure is causally related to the disease.[149] A number of formulae

have been derived by which the PEF can be estimated from epidemiological data. The formula giving an estimate from case–control or prospective studies is:

$$PEF = \frac{\pi.(RR-1)}{\pi.(RR-1)+1} \quad (1)$$

Here, *RR* is the relative risk of the disease in exposed groups relative to unexposed groups, and π is the prevalence of the exposure in the population of interest. Strictly *RR* should be an incidence ratio, or hazard ratio (HR), but the formula is a reasonable approximation where a relative risk is used, as long as incidence is low during the period under consideration. If the incidence is high enough to reduce the risk set by more than say 20%, the approximation will become poor.

There are several further points to note. First, the formula can be applied to case–control studies, substituting the odds ratio (OR) in place of the RR, if the study recruits an incidence density sample or if the disease is sufficiently rare for the OR to be a reasonable estimate of the RR. Second, the presence of π in the formula reminds us that the PEF has only a 'local' interpretation: it is not only a property of the disease and the exposure, but also of the time and place where the data were collected.

The most important caveat is that the formula is correct only when there are no confounding factors. Suppose, for example, we were considering the proportion of PID that is attributable to CT infection. We could estimate the OR in studies comparing CT prevalence in PID cases and non-PID control subjects. Typically, an OR of ≥ 6 is found in such studies (see *Chapter 7*). However, PID can be caused by other STIs as well as CT, and we can anticipate that women infected with CT are also more likely to be infected with other STIs as well. In the extreme case it would therefore be possible to observe a high PEF, even if there was no causal relation between CT and PID. Because of this, in the presence of positive confounders, we would have to consider the estimated PEF to be an upper bound. However, RRs adjusted for confounders can be used in Equation 1.

A simpler formula can be used in situations where there are no confounding variables:[148]

$$PEF = \frac{IP_{Popn} - IP_{UnExp}}{IP_{Popn}} \quad (2)$$

In this case IP_{Popn} and IP_{UnExp} are the cumulative proportions of the total population and the unexposed group developing the disease in a specified time period. (The comments above about rates and risks apply equally in this case.) Here the PEF can be interpreted as the proportional reduction in incidence in the population if the exposure was removed.

In the context of a randomised trial this corresponds to the idea of relative risk reduction (RRR), the proportion of incidence in the control group that can be eliminated by treatment:

$$RRR = \frac{IP_{Co} - IP_{Trt}}{IP_{Co}} \quad (3)$$

In CT screening trials (see *Chapter 7*), the outcome is incident PID in a 1-year period, the control group is the untreated population, and the treated group is screened and treated for CT. If we assume that the treatment is 100% effective, the RRR is therefore a trial-based estimate of the PEF.

Note, however, the importance of follow-up time in this calculation for these single screen studies. Screening and treatment for CT 'work' by preventing CT that is prevalent at the onset of the study from developing into PID during the 1-year follow-up. As the follow-up time is extended to 2, 3 or more years, women in *both* arms will experience more episodes of PID, and the RRR will converge towards zero. Further, this will be case whether the further PID episodes are CT related or not.

It is necessary, therefore, to consider models of PID incidence over a relatively long term, and compare scenarios in which (1) CT is a contributing cause of PID, and (2) CT does not exist. Estimates of the PEF taking the form of Equation 2 can be computed from these models of long-term effects.

Because we estimate a number of PEFs in the report, we use the notation PEF (risk factor←disease) to stand for the PEF of the disease attributable to the risk factor.

Statistical methods

Bayesian Markov chain Monte Carlo in WinBUGS

We have adopted Bayesian methods throughout this monograph.[150] Bayesian inference seeks to find the probability distribution of parameters (the posterior distribution) given the data (likelihood) and a prior distribution representing beliefs before the data were available. Classical inference (maximum likelihood) considers the parameter values at which the probability of obtaining the data is greatest. The Bayesian viewpoint is generally considered to be the more 'natural', in that people want to know what they can say about parameters given the data, not what they can say about the data given specific parameter values. However, it has been considered to have two disadvantages: the 'subjectivity' of the prior distribution and the computational difficulty of carrying out the high-dimensional integration required. We have avoided the subjectivity issue by using vague priors wherever possible (but see below). Advances in statistical theory[151] and computing power, and the development of highly flexible programming packages, such as WinBUGS,[44] have not only solved the computational problems, but also have made Bayesian methods the common choice for complex synthesis.[42]

The general principles of Bayesian MCMC are set out in standard texts,[152] and the procedures for Bayesian MCMC using WinBUGS are to be found in the WinBUGs on-line user manual.[44] The WinBUGS user is required to specify a model in algebraic terms. The model consists of (1) specification of prior distribution for 'basic'[50] parameters; (2) definitions of 'functional' parameters in terms of basic parameters, effectively the 'model'; and (3) specification of data likelihoods. Each item of data informs either a basic or a functional parameter.

Given some starting values for the basic parameters, parameter distributions converge to the joint posterior distribution. The WinBUGS 1.4.3 software makes use of a number of different MCMC algorithms. The most commonly used is Gibbs sampling,[153,154] in which samples are drawn from the conditional distribution of each parameter in turn, conditioning on the other parameters and on the data. Eventually, the marginal distributions of the parameters converge to their true marginal posterior distribution. In essence, all the parameters are uncertain: the marginal distribution expresses the uncertainty in each parameter, averaging over the distributions of all the other parameters.

Users find out about the posterior distributions of parameters of interest by sampling from them. Posterior means or medians are commonly cited, along with 95% CrIs or posterior standard deviations (SDs). CrIs have the interpretation that many people attribute to classical confidence intervals (CIs): namely that they express the probability that the parameter lies within that range, given the model, priors and data. Posterior SDs have the same interpretation as standard errors of estimates in classical statistics.

Because of the need for the process to converge, MCMC users must pay careful attention to a number of technical issues. First, diagnostics must be used to assess whether or not the process has converged. The Brooks–Gelman statistics[155] are commonly used as these are available in WinBUGS. The initial samples – the 'burn-in' – are discarded. Second, a large number of posterior samples may be needed to obtain accurate summary statistics. It is usually recommended that the process is run until the Monte Carlo error in the mean is $< 5\%$ of the posterior SD. It is also good practice to base posterior inference on more than one MCMC chain, because occasionally chains converge to different posteriors. Our practice here is to use two chains, to check they converge to the same point, to use a conservatively large burn-in period,

and also conservatively large numbers of posterior samples. In all the analyses in the monograph convergence has been achieved within about 25,000 samples, and we discard the first 50,000 as 'burn-in'. Posterior inference is based on 200,000 samples, 100,000 from each of two chains.

Sensitivity analyses

A further issue arising in Bayesian analysis is sensitivity to prior distributions.[150] Our general approach has been to introduce all information through likelihoods, and to put vague priors on all the basic parameters. When data are relatively sparse, or when the number of parameters to be estimated is large in relation to the amount of data informing them, results can be sensitive to the distributional form of the priors. Sensitivity analyses are therefore included wherever this is the case. The main findings from sensitivity analyses and their implications for the conclusions are discussed in the main text, but the full details are provided in the appendices.

In some cases it is preferable to fix parameters at a range of set values, to see the impact on the final results, rather than give the parameters a distribution. This, again, may generate a series of sensitivity analyses that are reported in full only in the appendices.

Model diagnostics and model selection

Our methods for assessing goodness of fit are based on Spiegelhalter *et al.*[156] The key statistic is the posterior mean residual deviance $\bar{D}$, corrected for the saturated model. This is the familiar residual deviance statistic from classical statistics, except that it is computed on each MCMC cycle rather than at the maximum likelihood values of the parameters. As a result a distribution of samples is generated and the mean is taken as a measure of goodness of fit. In a good-fitting model, the summed residual deviances should approximate to the number of data points.[156,157] When comparing models with different numbers of parameters a common approach in classical statistics is the Akaike Information criterion,[158] in which the models are 'penalised' by adding the number of parameters to the deviance statistic. We adopt this approach where appropriate using the Bayesian Deviance Information Criterion (DIC):[156,159] $DIC = \bar{D} + p_D$, where p_D represents the effective number of parameters. This statistic is calculated by WinBUGS and is available through the DIC tool. In models lacking any hierarchical elements, p_D is simply the number of free parameters in the model. With hierarchical models it reflects the degree of shrinkage. We make sparing use of hierarchical models in this monograph because of difficulties of interpretation (see *Chapter 4*), and they play no role in our final estimates.

$\bar{D}$ and DIC can be examined as measures of global model fit, or the contributions of individual data points to $\bar{D}$ and DIC can be examined in order to locate more precisely the particular data points that are in conflict with the other data under the model. It should be emphasised that these statistics are sensitive to prior distributions. This is one of the reasons why we tend to put vague priors on all parameters and introduce all of the information through data likelihoods. In this way, we can use residual deviance to assess conflict between different data sources.

There are exceptions to this rule, however, where we introduce evidence through informative priors. This is done where the parameter's role is local to a source of data, or when it is not central to the issues under study. In these cases, we may also use the WinBUGS 'Cut Function'.[44] This is a special function in WinBUGS that allows users to impose a prior distribution on a set of parameters that is not updated by the likelihood (the posterior is equal to the prior). The flow of information from elsewhere *back* to those parameters is blocked. This is done in cases when we wish to impose a belief (probability distribution) on a set of parameters, which we do not want to be altered by the data, or more pertinently, when we want the key parameters to be estimated conditional on a set of specific assumptions over which we have complete control.

An example that illustrates the reasoning behind this is the duration of symptomatic CT infection in women. It was difficult to source evidence for this parameter from formally collected data, and we turned, instead, to expert clinical opinion. The duration of symptomatic infection is the sum of the 'incubation

time' between infection and the appearance of symptoms; the time from symptoms until treatment is obtained; and the time from treatment to clearance of infection. It was considered by clinicians that a uniform distribution between 4 and 8 weeks would represent clinically informed belief about this parameter, and this was the chosen prior. This information was combined with information on incidence, prevalence, duration of asymptomatic infection, and the proportion of infections that are symptomatic, all of which had been based on formal consideration of likelihoods of data sourced in a reproducible way. If *all the* evidence was combined without the 'cut' function, the incidence and prevalence information would have changed our view of the duration of symptomatic infection.

In our view, this was not appropriate. We therefore used the 'cut' function to prevent this flow of information between sets of parameters. This allows us to estimate the other parameters, *conditional* on the assumption that this duration parameter is 4–8 weeks. This is clear and reproducible. If we built a full probability model with feedback, we would have to report that the parameter estimates were conditional on having a *prior* of 4–8 weeks on the duration parameter, but leading to a posterior distribution for duration that was somewhat different. We do, of course, assess the sensitivity of the results on incidence, prevalence, and other parameters to this assumption.

WinBUGS Development Interface and WinBUGS Differential Interface

WinBUGS itself has a limited repertoire of mathematical functions and statistical distributions. However, WBDev (WinBUGS Development Interface)[160] allows users to specify their own mathematical subroutines in the Component Pascal programming language. Although the functions must be evaluated on every MCMC iteration, the programming is 'hard wired' in a way that allows for very efficient computation. WBDev is used in to solve the discrete time Markov chain for repeat PID in *Chapter 8*.

A second WinBUGS 'add-on', WBDiff (WinBUGS Differential Interface)[161] allows users to specify systems of first-order ordinary differential equations, which are solved on each MCMC iteration using the Runge–Kutta method. WBDiff is used to solve the equations relating transition rates to transition probabilities in the continuous time Markov model fitted in *Chapter 6*.

When these add-ons are used, the WBDev and WBDiff code are provided in the appendices.

Directed acyclic graphs (influence diagrams)

The directed acyclic graph (DAG), or influence diagram, is a way of portraying the relationships between the data and the model parameters. The DAG is a formal entity in Bayesian statistics, and, in the WinBUGS language,[44] it is in fact possible to specify a model as a DAG rather than write WinBUGS code. In this monograph, following the previous work in Bayesian evidence synthesis,[42,50,52] we use schematic DAGs to show how all of the data sources 'fit together' around the synthesis model.

The schematic DAG in *Figure 5* illustrates how they are used and how they should be interpreted. The figure shows four numbered data items in rectangles, and five model parameters α, β, γ, δ, θ in ellipses.

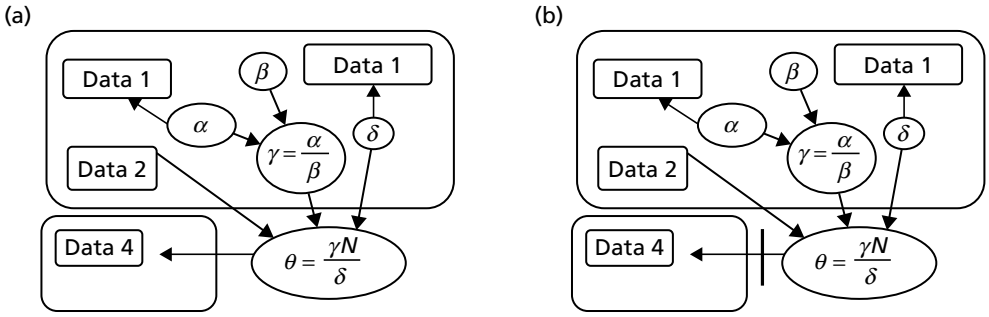

FIGURE 5 Directed acyclic graph. (a) Two sources of evidence on a parameter; and (b) one source is blocked from influencing the parameter.

The direction of 'influence' is shown in the direction of the arrows. The model is said to generate the data, so that the data are always pointed *to*. (In cases where information is introduced via informative priors, rather than as data, the information is shown in a rectangle, with an arrow pointing *to* the parameter.) Three of the five parameters, α, β, δ have no 'parents', that is, no arrows point to them, only from them. These are called 'basic' parameters,[50] and they are the parameters that must be assigned prior distributions. The remaining two parameters γ, θ are 'functional' or 'logical' parameters, and they can always be defined in terms of the basic parameters. The functional relationships between parameters are shown as equations in the ellipses, and these equations would be exactly those in the WinBUGs code.

Note that data can inform either basic or functional parameters. Also, the same model can be parameterised in many different ways, with different parameters defined as basic or functional. As long as the data are not too sparse in the relation to the number of parameters, this should not have any material impact on estimates.

Methods to examine evidence conflict

There are effectively two evidence groups in *Figure 5a*, each representing a different source of information on the parameter θ. One source of information is from Data item 4, the other is a combination of information from Data items 1, 2 and 3. *Figure 5a* represents a DAG in which both sources of information are combined to give a pooled estimate of θ. In this situation the inclusion of Data 4 not only informs θ, but it also feeds back to impact on the estimates of all the other parameters. We can examine conflict between the two groups of evidence by simply blocking Data item 4, to obtain an estimate of θ in its absence (see *Figure 5b*). Similarly, we can block out the other data to see what the estimate of θ would be if informed from Data item 4 alone. This technique is a version of 'node splitting'.[162,163]

A very useful way of assessing conflict between different sources of evidence visually is to plot the posterior densities. In this way, the posterior distributions of the parameter based on separate sources of evidence can be compared. The posterior density when the different sources are pooled can also be plotted on the same graph. This approach has been used by Dias *et al.*[164] in the context of splitting nodes in mixed treatment comparison evidence networks. A further refinement is to generate 'significance tests' based on Bayesian *p*-values: these test the hypothesis that one source of evidence generates a posterior distribution that is different from the other source.

Summary

1. Literature identification has been based on a high-yield strategy and snowballing, not on systematic review.
2. PEF is the proportion of incident cases in the population that would be eliminated by removing the exposure.
3. The synthesis models are fitted in a Bayesian MCMC framework using WinBUGS.
4. Global measures of goodness of fit, DIC and $\bar{D}$, are used to assist model selection.
5. Vague prior distributions are used throughout, so that conflict between evidence sources can be assessed by goodness of fit.
6. 'Node-splitting' methods are also used to assess conflict between different sources of data.

Chapter 4 Duration of asymptomatic *Chlamydia trachomatis* infection

Objective

To estimate the mean duration of asymptomatic infection in women, based on duration studies alone.

Introduction

There have been several attempts to estimate the duration of asymptomatic CT infection. It is intrinsically difficult to measure because, although one can determine whether an individual is infected by bacterial culture or DNA amplification, it is not possible to determine for how long the infection has been present. Further, infected individuals would now normally be treated, so it is not possible to observe the natural life course of the infection. Studies do exist, however, of untreated women. Previous reviewers have all noted that the published studies have produced a highly diverse set of estimates but have responded to the heterogeneity in different ways. Golden *et al.*[139] presented no summary estimate at all, whereas Geisler *et al.*[74] cited a median of approximately 1 year, based on an essentially narrative review. Korenromp *et al.*[142] derived an estimate of 1.4 years. This was based on an unweighted average of the study-specific estimates, in order to avoid giving a higher weight to larger studies.

Models of the transmission dynamics of CT have to make assumptions about its duration. Turner *et al.*,[141] citing the reviews by Korenromp *et al.*,[142] and Golden *et al.*,[139] drew somewhat different conclusion from the evidence, and assumed a 0.5-year duration. Roberts *et al.*[2] chose 0.56 years, citing the well-known modelling study of Kretschmar *et al.*,[137] which used the results from a duration study by Rahm *et al.*[165] More recently, Althaus *et al.*[166] based their estimate on a single study. It is worth emphasising that, although these authors have not used exactly the same evidence base, there is considerable agreement between them on the core studies that need to be included in a review.

In this chapter we compare fixed-effect (FE) and random-effects (RE) models for the duration of CT infection in a somewhat more formal way than previous authors, using statistical measures of goodness of fit. But we also re-examine the diverse study designs that constitute the evidence base. This leads us to propose a latent class analysis of the data, in which there are, in effect, two or three classes of infection, or apparent infection, each with a different expected duration. We assume that clearance rates within each class are constant over time, giving us a mixture of two or three exponentials.[167] Unlike the standard FE and RE models, which can be estimated relatively directly from the data, under the latent class models the study-specific proportions of individuals in each class represent an additional form of sampling variation. In the methods section (see below), we explain how the patient recruitment and sampling mechanisms engaged by each study design impact on the data likelihood. In the discussion section (see below), we consider other survival models and meta-regression models that could have been used, and briefly assess the scientific plausibility of the mixture-of-exponentials models we propose. The work in this chapter has been published.[168]

Methods

Evidence identification

Our target parameter is the mean duration of an episode of untreated, asymptomatic CT in women. Studies in any country on any population were included, as long as they reported sufficient information to determine: numbers at risk of clearing infection, numbers who cleared infection, and total time at risk. We define an episode of CT as the time from infection to the time that the infection has cleared, in the absence of treatment. Studies of CT duration have the inherent limitation that patients may clear infection and be re-infected during the follow-up period. For this reason we consider same-partner re-infections, which microbiological evidence suggest comprise the great majority of re-infections[169] to be part of a continuous episode. Symptomatic infection is defined as infection that is diagnosed and treated.

We identified studies through three recent reviews,[74,139,142] reference searches of published economic and population transmission models,[1,2,8,29,140,141,170–174] and consultation with an expert panel (see *Acknowledgements*). A 2012 search identifying papers that had cited the three reviews revealed no further studies of duration of asymptomatic infection in women. The recent review paper by Geisler *et al.*[74] was part of an authoritative review of CT epidemiology set up by an expert group of the CDC.[38] Two studies[175,176] were excluded because patients were treated for CT or other infection soon before recruitment[175] or during follow-up.[176] Three were excluded because they provided insufficient information.[177–179] Results from the 11 studies[73,79,165,169,180–186] included are shown below (see *Table 4*).

Initial consideration of the data: study design

The studies differed in their design with some being set in STD clinics, and others based on women identified as infected by population screening (*Table 4*). Two were set in antenatal clinics, where women were screened for CT at a prenatal visit. In most studies, patients were followed forward from an initial examination, when samples were collected, until they returned for a second visit, when CT status was assessed again; patients attended for the second test after different intervals but each patient was retested only once. In other studies,[169,180] the same patients were tested repeatedly over several follow-up periods: those at risk in subsequent periods had already 'survived' the previous period without clearing infection.

A preliminary analysis is presented in *Table 4*, showing the numbers clearing infection during each follow-up period i of study j with mean follow-up time t_{ij}. (Where this was not reported, we estimated it from the data at hand: details in *Appendix 1*.) We assume that in each study the clearance of infection in each woman is the result of an independent Bernoulli trial, and that the number clearing infection in each period is therefore binomially distributed with parameter $\theta(t_{ij})$, and a constant clearance rate λ_{ij}. As the observations are *interval censored*, the exact clearance times are unobserved, but we can obtain an estimate of the clearance rate λ_i and CIs through the relation $\theta(t_{ij}) = 1 - \exp(-\lambda_{ij}t_{ij})$.This is the complementary log–log (cloglog) link commonly assumed for binomial data generated in this way.[187] The mean CT duration Δ_{ij} in that segment is estimated by $\Delta_{ij} = 1/\lambda_{ij}$.

The crude estimates of mean duration vary by over 40-fold, varying between studies and also within. Although it is possible to fit either FE or RE models to these data, neither would capture the strong association between clearance rate, length of follow-up and study design. The studies based in STD clinics tend to have much shorter follow-up periods, and have a higher clearance rate than the longer-term studies, which tend to be screening studies. In the STD studies, women were recruited only if they were both asymptomatic at first attendance and had no indicators, such as a sexual partner with CT, that would trigger immediate treatment. Our interpretation is that women who attend STD clinics in the absence of any symptoms are most likely to be seeking to assess their infection status following a possible recent sexual exposure to CT. Those found to be infected in this setting might therefore have their first sample collected within a few days of the date of infection, and thus represent *incident* cases of infection.

In screening studies, observations are *left truncated* as well as interval censored: those found to be infected have already 'survived' clearance for an unknown period prior to screening. Screening studies

TABLE 4 Studies of the duration of asymptomatic CT infection in women: the study design, the numbers at risk, the numbers clearing infection, and the follow-up period over which they were observed.[168] Reproduced from: Mixture-of-exponentials models to explain heterogeneity in studies of the duration of *Chlamydia trachomatis* infection. *Stat Med* 2013;**32**:1547–60, under the Commons Attribution-NonCommercial license 3.0: http://creativecommons.org/licenses/by-nc/3.0/

First author	Study design	Number clearing CT	Number tested at follow-up	Follow-up period	Estimated mean follow-up, years	Mean duration, years (95% CI)
Johannisson[183]	Clinic	10	23	0–2w	0.038	0.07 (0.04 to 0.15)
		7	14	0–3w	0.058	0.08 (0.05 to 0.22)
		6	14	0–4w	0.077	0.14 (0.07 to 0.39)
		6	8	0-(5–8)w	0.125	0.09 (0.05 to 0.29)
Joyner[79]	Clinic	2	12	0-(2–7)d	0.012	0.07 (0.03 to 0.58)
		7	28	0-(8–14)d	0.030	0.10 (0.06 to 0.27)
		1	4	0-(15–21)d	0.049	0.17 (0.05 to 7.72)
		0	8	0-(22–42)d	0.088	
		3	6	0-(43–231)d	0.274	0.39 (0.18 to 2.14)
Geisler[73]	Clinic	23	129	0-(4–59)d	0.045	0.23 (0.16 to 0.36)
Paavonen[184]	Clinic	3	15	1m	0.083	0.36 (0.15 to 1.80)
Rahm[165]	Screened	17	85	3m	0.250	1.12 (0.74 to 1.93)
		0	1	6m	0.500	
		0	1	9m	0.750	
Sorensen[185]	Screened	8	13	(2–24)m	1.000	1.05 (0.6 to 2.62)
McCormack[186]	Screened	3	7	(16–17)m	1.375	2.46 (1.11 to 13.3)
Morre[180]	Screened	2	20	1m	0.083	0.79 (0.29 to 6.66)
		2	5	6m	0.500	0.98 (0.4 to 9.21)
		4	15	1 + 5m	0.417	1.34 (0.63 to 5.12)
		0	1	1 + 11m	0.917	
		2	13	6 + 6m	0.500	2.99 (1.12 to 25.4)
Molano[169]	Screened	44	82	1y	1.000	1.32 (1.00 to 1.86)
		23	37	1 + 1y	1.000	1.03 (0.72 to 1.68)
		7	14	2 + 1y	1.000	1.44 (0.80 to 3.81)
		2	6	3 + 1y	1.000	2.47 (0.97 to 22.6)
Sheffield[181]	Antenatal	16	52	(1–5w)	0.058	0.16 (0.10 to 0.28)
		45	88	(6–13w)	0.183	0.24 (0.18 to 0.34)
Alexander[182]	Antenatal	12	15	6w	0.115	0.07 (0.05 to 0.16)

d, days; m, months; w, weeks; y, years.
Follow-up time: 0–Xw indicates follow-up for X weeks of women who were observed to be infected at time zero; 0–(X–Y)d indicates between X and Y days' follow-up from time zero; 3 m indicates a 3-month period of observation with an unknown starting date; 1 + 1y indicates a 1-year period of observation of women who had already been followed for 1 year without having cleared infection. CrIs based on the binomial distribution (see text).

therefore reflect *prevalent* infection. This, together with the apparently slower clearance rate seen in screening studies, provides a prima facie case for believing that CT infections may clear at different rates: screening studies may generate apparently longer clearance rates because they selectively oversample the longer-lasting infections, as short-duration infections are more likely to clear before screening, a phenomenon known as length-biased sampling.[188]

The high rates of clearance reported at the shortest follow-up times are consistent with observations made in *Chapter 2* on the likelihood of 'passive infection' (see also Joyner *et al.*[79]), according to which CT is recoverable from the female genital tract for several days after sexual intercourse, but without leading to an active infection.

This provides a basis, independent of the evidence in *Table 4*, for a two-class model comprising passive and true infections. However, examination of *Table 4* suggests that, even within the screening studies, longer follow-up times seem to be associated with lower average clearance rates. We have therefore also considered a three-class model, comprising passive infection, and fast- and slow-clearing infections. These alternative conceptions of the natural history of transmission, infection and clearance can be captured by mixture-of-exponentials models. We return to a consideration of the *a priori* scientific plausibility of both the two- and three-class models in the discussion, where the role of alternative survival models is also considered.

We begin by describing the standard FE and RE models for these data. We then describe the parameters of the mixture of exponentials models and develop expressions for their functional relationship to the parameters in the likelihood.

Statistical models for studies of the duration of Chlamydia trachomatis *infection*

Random-effects model

The RE model assumes a single clearance rate λ_j within each study j, drawn from a distribution of clearance rates. We assume that the log study-specific rates are normally distributed, with a mean μ and variance σ^2 to be estimated from the data: $\log(\lambda_j) = \alpha_j$, $\alpha_j \sim N(\mu, \sigma^2)$.

In studies reporting the numbers clearing infection in more than one period of follow-up, the same rate applies in each time period, so that the number clearing infection during the ith period in study j is binomially distributed: $r_{ij} \sim Bin(\theta(t_{ij}), n_{ij})$, where $\theta(t_{ij}) = 1 - \exp(-\lambda_j t_{ij})$.

Fixed-effect model

The FE model follows immediately from setting $\lambda_j = \lambda$, for all j.

Mixture of exponentials model

In both FE and RE models described above, within every study, each woman has the same clearance rate. We now develop the three-class mixture model. This allows women within every study to have one of three different rates. It is the same three rates that are present in every study. We then introduce the two-class model as a special case. As we now show, there are three types of data: each provides information on a different function of a common set of parameters. For each type of data we define the likelihood and explain the relation between the parameters in the likelihood and the common set of rate and probability parameters in the model.

We assume that *incident infections* are in one of three categories ('passive', fast-clearing and slow-clearing). The proportion of incident infections in each of the three classes is p_1, p_2, p_3; $p_3 = 1 - p_1 - p_2$, and the clearance rates are λ_1, λ_2, λ_3, with each rate having a fixed effect across studies. It can readily be established that the two-class model is a special case of the three-class model, in which $p_3 = 0$, and that the FE model is a one-class model in which $p_2 = 0$ and $p_3 = 0$. We henceforth refer to the latter as the one-class FE model. The schematic DAG (*Figure 6*) clarifies the relationships between the shared rate and

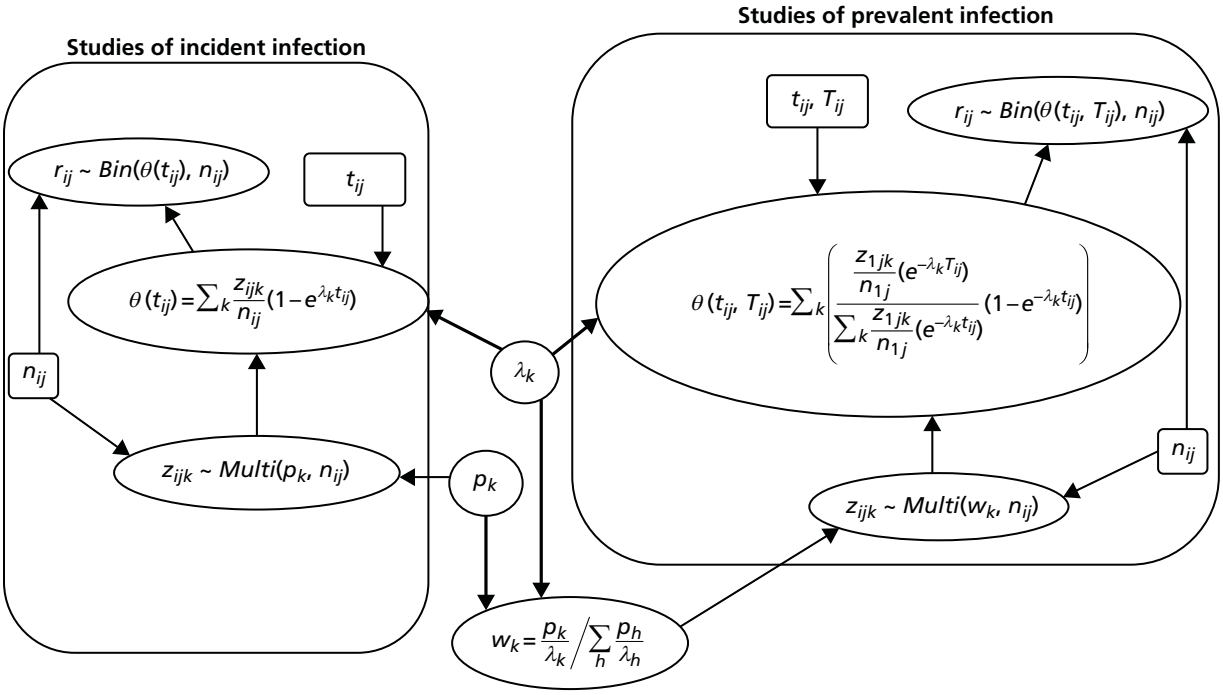

FIGURE 6 Directed acyclic graph of the CT duration synthesis. Subscripts: *i*, observation; *j*, study; *k*, class; *n*, number followed up; p_k, proportion of infections in class *k*; *r*, number of patients clearing infection; t_{ij}, duration of follow-up period; T_{ij}, duration of previous segments; w_k, expected proportion of infections in class *k*; z_{ijk}, number of infections in class *k* at risk of clearing in time interval *ij*; $\theta()$, probability of clearance; λ_k, clearance rate of class *k* infections.

proportion parameters, and the likelihoods generated by the sampling and recruitments schemes represented by the three study designs, under the mixture-of-exponentials models.

Sexually transmitted disease clinic studies

We interpret clinic studies as studies of incident infection, as explained above, and the infections observed are therefore a mixture of the three types of infections, the clearance rates of which are exponentially distributed. In any given sample, the denominator number of infections 'at risk' of clearing is an unobserved mixture of infection types. The proportions recruited from each class *k* will vary by chance around the population mean proportions p_1, p_2, p_3. For each observation *i* in study *j*, the number of infections $z_{i,j,k}$ in class *k* at risk of clearing will be drawn from a multinomial distribution, with probabilities p_1, p_2, p_3 and denominator n_{ij}: $z_{i,j,k=1,2,3} \sim$ Multinomial (p_1, p_2, p_3, n_{ij}).

The expected clearance probability in follow-up period t_{ij} is therefore a weighted average of the clearance probabilities for the three classes with the z_{ijk} as weights:

$$\theta(t_{ij}) = \frac{z_{i,j,1}}{n_{ij}}.\left(1 - \exp(-\lambda_1.t_{ij})\right) + \frac{z_{i,j,2}}{n_{ij}}.\left(1 - \exp(-\lambda_2.t_{ij})\right) + \frac{z_{i,j,3}}{n_{ij}}.\left(1 - \exp(-\lambda_3.t_{ij})\right) \quad (4)$$

Screening studies: follow-up from the initial screen

The expected proportion of incident infections of each type is p_1, p_2, p_3. But screening studies must over-sample longer lasting infections. The prevalence π_k of infection of type *k* can be derived from the relationship: *prevalence* = *incidence* × *duration*:

$$\pi_k = \frac{\eta.p_k}{\lambda_k} \quad (5)$$

where η is the incidence of CT. Thus the expected proportion w_k of women in each group k in a prevalent population is:

$$w_k = \left(\frac{p_k}{\lambda_k}\right) / \left(\frac{p_1}{\lambda_1} + \frac{p_2}{\lambda_2} + \frac{p_3}{\lambda_3}\right) \tag{6}$$

Again, to take account of the chance variation between the proportions in the general population and the proportions sampled for each observed group, we take the data collected in each observation as a multinomial sample, now with parameters w_k. For the *first* observation ($i = 1$) in study j we may now use the same expression (see equation 4) for the probability of clearance, with weights:

$$Z_{1,j,k=1,2,3} \sim Multinomial(w_1, w_2, w_3, n_{ij}) \tag{7}$$

Screening studies: subsequent follow-up periods

In the Molano *et al.*[169] and Morre *et al.*[180] studies, after the first observation period, patients who are 'at risk' of clearance in the subsequent follow-up segments should be considered to have 'survived' both the original period between infection and screening, and the previous follow-up segments, of duration say T_{ij}, since screening. This is shown in *Table 4*. For example in the Molano *et al.*[169] study, 2 out of 6 cleared CT during a 1-year period of observation, having already 'survived' a further $T_{ij} = 3$ years since their recruitment to the study. This leads to another, rather complex, sampling structure. The proportion $z_{i,j,k}$ in each category k for each observation i in study j is determined by two factors. First, a single multinomial sample at the start of the first time period (the time of screening) determines the numbers w_k in each group, exactly as in equation 6. Second, in subsequent time periods there is no further multinomial sampling. Instead, each original sample in category k must be depleted by a further $\exp(-\lambda_k T_{ij})$, so that, through substitution into equation 5, and then equation 4, the observed number of women clearing CT in each time period T_{ij} is equal to:

$$\theta_{ij}(t_{ij}, T_{ij}) = \left(\sum_{k=1}^{3} \frac{\frac{z_{1,j,k}}{n_{1,j}} \exp(-\lambda_k T_{ij})}{\sum_{k=1}^{3} \frac{z_{1,j,k}}{n_{1,j}} \exp(-\lambda_k T_{ij})} (1 - \exp(-\lambda_k t_{ij})) \right) \tag{8}$$

Studies based in antenatal clinics

Two studies[181,182] were set in antenatal clinics. One clinic[181] recruited pregnant women between 8 and 23 weeks of gestation, and the other clinic[182] between 36 and 40 weeks of gestation. These studies, in principle, could be considered as having the same recruitment process as a screening study. However, it would also be plausible to consider at least a proportion of those recruited to have been infected at the same time that they conceived, on average, say, 16 weeks or 38 weeks prior to the initial sample. We therefore explored including the antenatal studies under the extreme assumptions that (1) they are left-truncated screening studies, and (2) infection was acquired at the point of conception.

Sensitivity and specificity of diagnostic tests for Chlamydia trachomatis *used in the studies*

Poor test sensitivity may overestimate the clearance rate, as patients who have not cleared infection may be thought to have done so. Most of the studies use NAATs, which we assumed to be 100% sensitive and 100% specific. Sensitivity is defined as the probability of detecting a CT infection, given that it is present. For example, if sensitivity of culture is ψ then studies of clearance rate that test for CT using culture estimate $\theta_{ij}(t_{ij})/\psi$, rather than $\theta_i(t_i)$. Geisler *et al.*[73] used both NAAT and culture tests, and reported the sensitivity of culture given a previous culture positive result for that infection as 78 out of 86 (90%), and we use this as an estimate of ψ.

Note that we distinguish between test sensitivity in the initial testing to determine infection status, and sensitivity in subsequent tests to establish whether or not the infection has cleared. A failure to detect infection initially will mean that the individual is not recruited into the study, and we assume that this has no bearing on the estimates of clearance rates. Although we can estimate the sensitivity of culture in women who are initially culture positive, it is difficult to estimate sensitivity of NAAT in women who are initially NAAT positive, and we make the simplifying assumption that this is 100%. An alternative, but equivalent, assumption would be that if the bacterial load is so low that it cannot be detected by NAAT, then it does not constitute a clinically relevant infection.

Prior distributions

For the RE model, we place a flat normal prior on the log mean clearance rate and a *Uniform*(0,5) prior on the log SD, and the effect of this prior is assessed through sensitivity analysis. Because the shortest follow-up period is 1 week, the *Table 4* data do not allow us to adequately identify λ_1, the rate of clearance of passive infection, in the three-class model. To improve the identifiability, we set $\lambda_1 = 120$ per year, implying that passive infection lasts, on average, only 3 days. We confirmed subsequently that the precise value chosen has negligible impact on the other parameters. We assigned priors of the following form to the clearance rates: $\log(\lambda_2) \sim Uniform(A,B)$, and $\log(\lambda_3) \sim Uniform(C,\lambda_2)$. We chose the values $A = -1$, $B = 4.8$ and $C = -2.62$, corresponding to durations between 3 days and 2.7 years for the duration of a fast-clearing infection and between $1/\lambda_2$ and 13.7 years for slow clearing. As a sensitivity analysis for the three-class model, we varied the values of A, B and C, and also considered placing Uniform priors on clearance rate, and on duration (1/clearance rate), as alternatives. We also estimated a two-rate model, in which $p_3 = 0$, $\lambda_3 = 0$, and $\log(\lambda_3) \sim Exponential(0.001)$. A prior distribution for the sensitivity of diagnostic tests of $\psi \sim Beta(78,8)$ was based on the results from Geisler *et al.*[73] The mixing proportions p_1, p_2, p_3 were given *Beta*(1,1) and *Dirichlet*(1,1,1) priors in the two- and three-rate mixture models, respectively.

Estimation of Chlamydia trachomatis *infection duration from each model*

The target parameter is the average duration of asymptomatic infection. For the one-class FE model, this is simply $1/\lambda_1$. For the RE model, at least two estimates can be generated, depending on how the model is interpreted (see *Discussion*, below). One would focus on the RE mean, which would generate an estimate of duration. In this case we monitor a node $1/\exp(\mu)$ and report its summary statistics. Another interpretation might be that the rate obtained in any future situation is effectively another sample from the RE distribution whose parameters have been estimated from the data. The estimate produced by this 'predictive' interpretation is readily found in the MCMC framework, by monitoring the nodes $\alpha_{new} \sim N(\mu,\sigma^2)$ and $1/\exp(\alpha_{new})$, and reporting summary statistics for the distribution of the latter.

In the two-class mixture model, a proportion p_1 of CT+ patients do not have a true infection, so the estimate of CT duration is $1/\lambda_2$. Under the three-class mixture model, the estimate of CT duration is a weighted average of the durations of the fast- and slow-clearing infections:

$$\frac{p_2}{\lambda_2(p_2+p_3)} + \frac{p_3}{\lambda_3(p_2+p_3)} \tag{9}$$

The WinBUGS code for the two-class model is presented in *Appendix 2*.

Results

Model fit statistics

There were a total of 28 data points from 11 studies informing CT duration, constituting the independent observations for the evidence synthesis. The models that include the antenatal data assuming it represents prevalent infection fit better than when these data are assumed to represent infection that is incident at the time of conception (results not shown). The model fit statistics (*Table 5*) show, however, that when the antenatal data (assumed to be prevalent cases) is included *none* of the models fits well.

TABLE 5 Model fit statistics: one-class FE, RE, two-class and three-class models.[168] Reproduced from: Mixture-of-exponentials models to explain heterogeneity in studies of the duration of *Chlamydia trachomatis* infection. *Stat Med* 2013;**32**:1547–60

Model fit statistics	With antenatal studies, 28 data points				Without antenatal studies, 25 data points			
	One-class, FE	Two-class	Three-class	RE	One-class, FE	Two-class	Three-class	RE
D-bar	268	128	76	36	170	25	25	31
p_D	1.6	7.5	6.8	10.7	1.4	6.8	6.9	9.8
DIC	270	136	82	47	171	32	32	40

D-bar residual deviance; DIC, Deviance Information Criterion; p_D effective number of parameter.

The residual deviance $\bar{D}$ is expected to approximate the number of data points (28), but the best-fitting model, the RE model, has a $\bar{D}$ of 36. The FE and the latent class models fit even worse. If the antenatal studies are excluded, the two- and three-class models now fit the data well. The RE model is an acceptable fit, but the DIC, a measure that sums $\bar{D}$ and the number of effective parameters p_D, strongly favours the mixture models, as they have fewer effective parameters as well as having a better fit. Further results are reported for only the reduced data set, excluding the antenatal studies, to which we return in the discussion.

Parameter estimates

The rate estimates from each model are shown in *Table 6*. The rate estimate generated by the one-class FE model (1.1 per year) is lower than the RE model (1.7 per year) because the former is more strongly influenced by the largest studies, which happen to be the screening studies with their longer follow-up and

TABLE 6 Posterior summaries – mean (2.5th, 50th, 97.5th) percentiles – of parameter estimates.[168] Reproduced from: Mixture-of-exponentials models to explain heterogeneity in studies of the duration of *Chlamydia trachomatis* infection. *Stat Med* 2013;**32**:1547–60

Parameter		Symbol	One-class FE	Two-class	Three-class	RE
Clearance rate	Class 1	λ_1	1.09 (0.93, 1.09, 1.26)	120[a]	120[a]	–
	Class 2	λ_2	–	0.74 (0.61, 0.74, 0.89)	9.5 (0.69, 1.77, 74.7)	–
	Class 3	λ_3	–	–	0.64 (0.25, 0.66, 0.84)	–
Proportion	Class 1	p_1	1	0.23 (0.16, 0.23, 0.31)	0.19 (0.03, 0.20, 0.29)	–
	Class 2	p_2	–	0.77 (0.69, 0.77, 0.84)	0.31 (0.01, 0.26, 0.76)	–
	Class 3	p_3	–	–	0.49 (0.02, 0.56, 0.79)	–
Mean clearance rate		$\exp(\mu)$	–	–	–	1.70 (0.57, 1.51, 3.89)
Predictive clearance rate		$\exp(\alpha_{new})$	–	–	–	18.5 (0.08, 1.50, 28.9)
Between-study SD		σ	–	–	–	1.33 (0.74, 1.24, 2.46)

a Note that λ_1 was fixed at 120 (see text).
One-class FE, RE, two-class and three-class models. Rates are per year.

slower clearance, whereas the latter allows greater influence of the more numerous high-clearance-rate studies. A prediction for a new study has an expected rate of 18.5. This is due to the skew in the posterior distribution (median 1.5). The degree of between-studies heterogeneity according to the RE model, as indicated by the between-studies SD, is extraordinary. The posterior mean of the between studies SD is 1.33 on the log scale, indicating that a study on the 95th percentile of clearance rate would have a clearance rate almost 200 times higher than the fifth percentile. The mean clearance rates estimated from the exponential mixture models were 0.74 per year for the two-class model, and 9.5 and 0.64 per year for fast and slow clearers, respectively, in the three-class model. They identify 19–23% of infections as 'passive', clearing exceptionally rapidly. Note that the three-class mixture model is barely identified by the data, with the proportions in each class and the class-specific clearance rates having extremely wide CrIs. This is largely a result of the sparseness of data at long follow-up times.

The estimates of the target parameter, mean duration of CT (*Table 7*), indicate broad agreement between the two- and three-class models (1.36 years, 1.24 years, respectively), and considerably lower estimates from the one-class FE and RE models. The posterior medians of the mean duration in the RE model, whether based on the mean or on the predictive distribution, are barely half those from the mixture models. The RE distribution is so skewed, however, and its variance so large, that the mean of the predictive distribution is 23 years.

The 'observed' durations in each observation from *Table 4* are compared with those predicted by the two-class model taking into account study design and duration of follow-up, and with duration plotted on a log scale (*Figure 7*). Note that for the longer-duration screening studies the predicted duration stabilises to a single value, which is, in fact, the expected duration of true CT infection. This is because in a screening study the chances of recruiting a woman with passive infection are very low indeed. This is also evident from equation 6. It is interesting that the results from the longer-duration screening studies systematically deviate from the two-rate model in showing that observed duration increases with length of follow-up. This is evidence that favours a three-rate model but, because the data are so sparse, the global statistical fit does not support the three-rate against the two-rate model.

Sensitivity analysis

Sensitivity to priors was also explored. When the RE variance prior was changed from Uniform(0,5) to Uniform(0,3), this changed the posterior expectation of the RE mean by less than 1%. Similarly, the posterior between-studies SD was changed by only 1%. This suggests the prior is having only a small effect on the between-study variance and we conclude that the *Table 4* data are adequate to estimate the RE model.

TABLE 7 Posterior summaries – mean (2.5th, 50th, 97.5th) percentiles – of the mean duration of CT infection in years.[168] Reproduced from: Mixture-of-exponentials models to explain heterogeneity in studies of the duration of *Chlamydia trachomatis* infection. *Stat Med* 2013;**32**:1547–60

Model	Expression for mean CT duration, under each model	Mean (2.5th, 50th, 97.5th) percentiles
One-class FE	$1/\lambda_1$	0.92 (0.79, 0.92, 1.07)
Two-class	$1/\lambda_2$	1.36 (1.13, 1.35, 1.63)
Three-class	$\frac{p_2}{\lambda_2(p_2+p_3)}+\frac{p_3}{\lambda_3(p_2+p_3)}$	1.24 (0.85, 1.25, 1.59)
RE (mean)	$1/\exp(\mu)$	0.75 (0.26, 0.66, 1.76)
RE (predictive)	$1/\exp(\alpha_{new})$	23.2 (0.03, 0.67, 13.24)

p_k, proportion of infections that are type k; α_{new}; log clearance rate in a new study under the RE model; λ_k, clearance rate in infections of type k; μ, population mean log clearance rate in a RE model.
One-class FE, RE, two-class and three-class models.

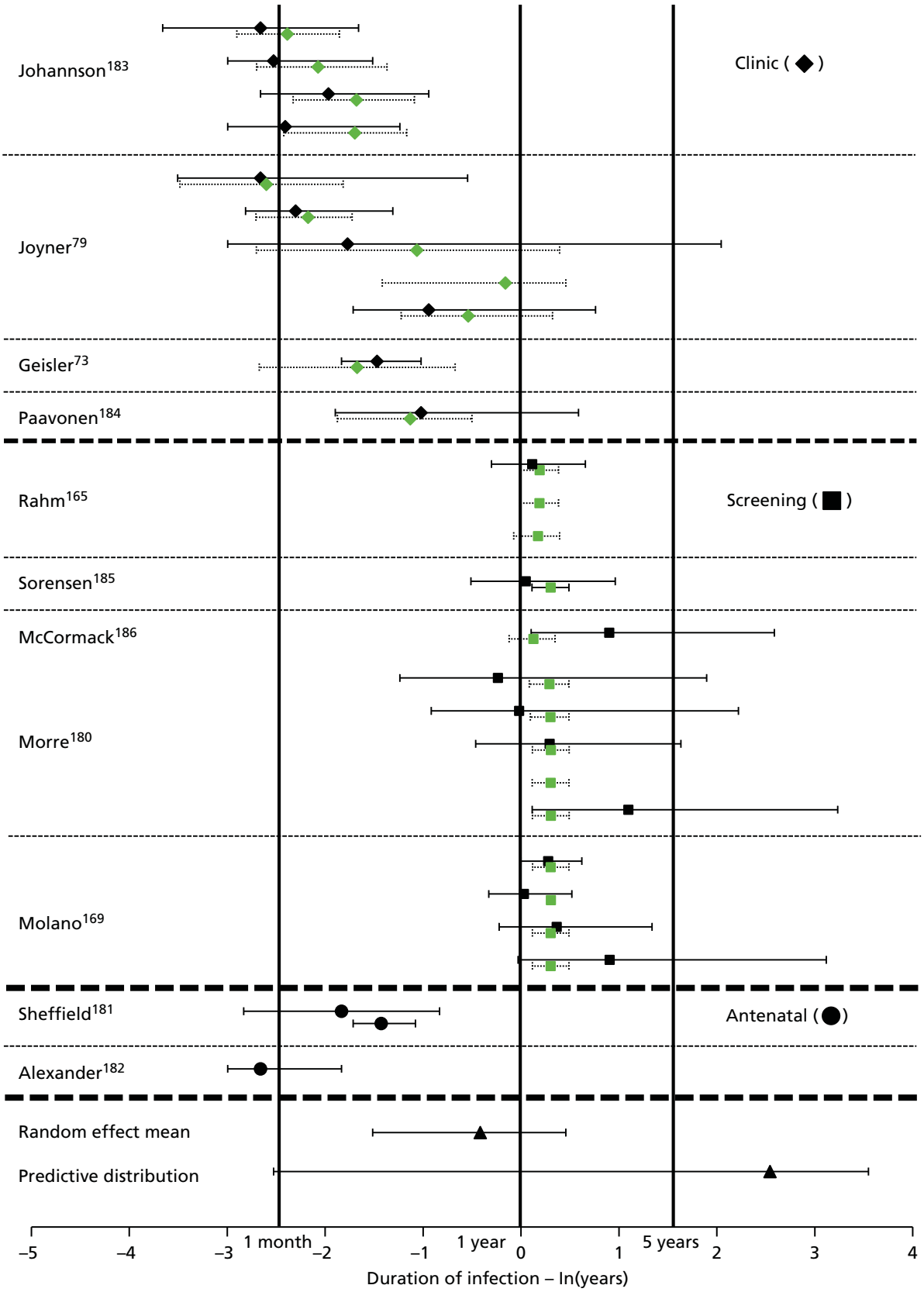

FIGURE 7 Forest plot of the 'observed' log duration (black) and predicted (green) based on the two-rate model, mean and 95% CrI, for each study observation (in the same order as *Table 4*).[68] Adapted from: Mixture-of-exponentials models to explain heterogeneity in studies of the duration of *Chlamydia trachomatis* infection. *Stat Med* 2013;**32**:1547–60. Results of the RE meta-analysis are presented at the base.

We examined a two-class model in which λ_1 was not restricted to a fixed value: it did not fit better and has the same residual deviance as the model restricted to $\lambda_1 = 120$. It converged to a very high rate of clearance of passive infection, with p_1 estimated at 0.23 (0.16,0.32), very close to the estimate from the restricted model. This finding supports our use of the value of 120 to improve model stability. All reasonable priors for λ_2 gave essentially identical answers. Sensitivity analysis of the three-class models showed that the rate and probability parameters were sensitive to priors, which might be expected from the fact that the model is only barely identifiable. The estimated duration of CT, the target parameter, was, however, relatively insensitive, with values ranging from 1.17 to 1.24 years.

Discussion

We have compared standard fixed- and random-effect models with mixtures of exponentials models,[167] in which the proportions are latent variables estimated from the data. The latter are considerably more complex to fit because of the impact of the patient recruitment processes in the different study designs. This had two aspects: first, once one abandons models in which there is a single constant clearance rate, it is necessary to take into account the left truncation in the screening studies.[188] Second, although the latent class model proposes that the proportions of incident infections belonging to each class is a fixed property of chlamydia infections, each different study in effect constitutes a different sample in which the proportions in each class differ from their expectations owing to both sampling variation and selective attrition in previous periods.

The choice between the alternative models must be considered both from the viewpoint of model fit and in terms of scientific interpretation. In this application, we found that the latent mixture models both provided a better fit and used fewer parameters, but this was true only when the two antenatal studies[181,182] were excluded. With the antenatal studies[181,182] included, none of the models fitted particularly well, although the RE model gave by far the best fit. These studies have unusually high clearance rates, considering they are not studies of incident infection, and this is at variance with what would be expected biologically.[28,189] However, the higher clearance rate in Sheffield *et al.*[181] may have been because of problems with the ligase chain reaction (LCR) NAAT test, which was withdrawn in 2001 owing to problems including non-repeating positives.[190] The other antenatal study, Alexander *et al.*,[182] recruited patients while pregnant, who may be less likely to resolve infection than non-pregnant women, and retested them 6 weeks post partum when their cell-mediated immune response was likely to have returned to normal. Clearance rates would thus be expected to be higher than for non-pregnant women. This is, however, post-hoc reasoning.

The two-class model not only fits the data very well with a very small number of parameters, but it also represents a biologically driven model of transmission and infection (see *Chapter 2*). Interestingly, the concept of 'passive infection' was also proposed by the investigators of one of the duration studies.[79] The further distinction between Fast and Slow clearers is attractive, as it could reflect potential differences in the human protective immune response to CT,[74,191] but there is no previous suggestion that there is a relation between clearance rates and immune response to CT. Although there is little prior support for the three-class model, it is important to note that the target parameter, mean duration of CT, is not particularly sensitive to choice of two- or three-class mixture model, from which we conclude that results are relatively insensitive to model assumptions.

The analysis provided by the two- and three-class models is dependent on our conception of the studies as clinic-based or screening studies, and on our assumption that clinic studies recruit women with incident infection, while screening studies recruit women with prevalent infection. This is a somewhat simplistic view of the clinic studies, where women may be attending at varying intervals from the time of transmission, but it is required to make the analysis tractable.

None of the models examined allowed for re-infection. However, the model including fast and slow clearers provided no improvement in residual deviance, and the bias introduced by not accounting for CT re-infection in duration studies is far lower than the bias introduced by not accounting for clearance in studies of incidence. This conclusion is supported by a recent analysis by Althaus *et al.*,[166] who fit a model to data extracted from Kaplan–Meier curve in the Molano *et al.*[169] study. They showed that a single-rate model provided a good fit, and found that allowing for re-infection had almost no impact on estimates of duration.

Setting aside their generally poor fit, the RE models are not scientifically plausible because they imply an extraordinary degree of between-study variation in clearance rates, a 300-fold difference between 2.5th and 97.5th percentile. This seems to defy any biological interpretation.

Among the further models that could be advanced would be a meta-regression against length of follow-up, or a subgroup analysis by study design, perhaps with RE within groups. We have not fitted such models on the grounds that they cannot be used to provide an estimate of the true duration of CT infection for modelling purposes. This is because they assume that the duration of CT infection is not a biological property of CT, but depends on the duration and/or the design of the studies carried out to estimate it. Even if these models could be shown to fit the data, they would, at best, be no more than descriptively adequate: they lack explanatory power as well as public health utility.

Another model that could be suggested would have a two-rate latent class model, similar to the one we have proposed, but with random between-studies variation in the clearance rates in each class. This model has a DIC of 33: one point higher than the two- and three-rate models. The model is certainly a possible view of the data, but random variation between studies again creates problems of interpretation when applied to a public health context, with an estimate based on the RE mean (1.50 years, 95% CI 0.75 to 2.68 years) differing from an estimate based on the predictive distribution (27 years, 95% CI 0.29 to 6.7 years, median 1.4 years).

Latent class models are not the only models possible. A wide range of failure time distributions, such as Weibull or Gamma, in which the clearance rates can decline over time, could also have been fitted to these data, although biological interpretations of such models would perhaps be less compelling. In this connection it is worth noting that Weibull distributions with shape parameter less than 1 can always be generated by mixtures of exponentials.[167]

Asymptomatic infection: summary of findings and assumptions

Assumptions

1. Studies of duration have been either based in STI clinics, or they have been studies following individuals whose infection was detected by screening. We have assumed that these two designs recruit incident and prevalent infections respectively.

Summary of findings

1. The data are not compatible with a single-class (FE) model, and the degree of variability implied by the RE model is not scientifically plausible.
2. The exponential mixture models, whether two- or three-class, fit the data best, and accord with our understanding on the biological mechanisms of transmission and infection. The data does not distinguish between two- and three-class models.
3. The expected duration of asymptomatic infection in women, based on this data alone, is 1.36 years (95% CI 1.13 to 1.63 years), based on the two-class model.

Chapter 5 Synthesis of data on incidence, prevalence and duration

Objectives

To:

1. determine whether the estimated duration of asymptomatic CT infection in women, developed in the previous chapter, is consistent with the available evidence on incidence and prevalence of CT in the UK
2. produce a coherent set of estimates of incidence and prevalence by age group, and duration, and the proportion of infections that are asymptomatic.

Introduction

This chapter sets out to produce a set of age group-specific estimates of CT incidence and prevalence in the general population of women in England. First, we re-analyse the data from the LaMontagne *et al.*[192] incidence study accounting for CT clearance, informing clearance rate from the *Chapter 4* synthesis of the duration of asymptomatic infection. Second, we use information on setting-specific prevalence[106] of CT in the UK – which included estimates in the general population as well as in GP, FP and STD settings – to 'recalibrate' the estimated incidence rates from LaMontagne *et al.*[192] to the general population setting. Third, we exploit the well-known epidemiological relationship: *prevalence* = *incidence* × *duration* to generate an independent set of incidence estimates based on prevalence and duration data. This provides a degree of independent validation for the estimates obtained directly from the incidence study. Finally, having established there is consistency between the alternative estimates, we produce a coherent set of estimates of age-specific incidence and prevalence, and duration, in women in the general population that both conform to the appropriate epidemiological relationships and are based on all the available data. The work in this chapter has been published.[193]

Methods

Evidence sources

Duration of asymptomatic infection in women, and proportion symptomatic

For the duration of asymptomatic CT infections, Δ_A, we use the estimate of 1.36 (95% CI 1.11 to 1.62) years based on the two-rate model described in *Chapter 4*. The proportion of CT infections, φ, in which symptoms develop can be estimated from studies where asymptomatic women within a few days of exposure are followed without treatment to determine if symptoms develop. The one study[194] of this type reported that 26 out of a total 115 women developed symptoms, estimating φ at 23% (95% CrI 16% to 31%).

Duration of symptomatic infection in women

We define the duration of symptomatic infection, Δ_S, as the time between the point at which the patient becomes infected, and the point at which the infection is diagnosed, or the patient is empirically treated and the infection is cleared. This could be derived from information on the incubation period of CT and studies of time taken to seek health care in women who are subsequently diagnosed with CT. A recent literature search[142] found no data on incubation period, and although there were studies of time to seek health care in women with genital symptoms, specific information on those diagnosed with CT

was not found. Based on a clinical understanding of incubation time, and the time taken to seek and obtain treatment, we placed an informative prior on the time from infection to diagnosis, assuming it is uniformly distributed between 4 and 8 weeks, and, that, once diagnosed, a woman would not participate in a prevalence survey. We assess sensitivity to this by fitting a model where the duration of symptomatic infection varies uniformly from 3 to 12 weeks. The distributions used for each duration parameter are summarised in *Table 8*.

Chlamydia trachomatis incidence data

Only one published report on incidence in England was identified: LaMontagne *et al.*[192] Women aged 16–24 years were screened for chlamydia in GP, FP and STD clinic settings in two areas in England, and were followed prospectively at 6-month intervals for 6–18 months to assess CT infection and re-infection. However, this study is restricted to GP and clinic patients and does not address incidence in the English general population. A LCR test was used, for which we assumed 100% sensitivity and specificity. *Table 9* gives the proportions of 6-month-long observations in which CT-negative (CT–) women were CT+ at follow-up. These are divided into 'infections' and 're-infections', the latter being infections observed in women who were CT+ on recruitment or were infected for the second time during the follow-up period. The data are reported for age groups ($a = 1$, 16–17 years; $a = 2$, 18–20 years; $a = 3$, 21–24 years).

TABLE 8 Data on duration (years) of CT Infection, and proportion symptomatic. From Price *et al.*[193]

Parameter	Mean (95% CrI)	Source
Duration of asymptomatic infection	1.36 (1.11 to 1.62)	see *Chapter 4*
Duration of symptomatic infection	0.115 (0.079 to 0.151)	See text
Duration of symptomatic infection (sensitivity analysis)	0.144 (0.062 to 0.227)	
Proportion of incident infections in which symptoms develop	0.231 (0.159 to 0.311)	Geisler 2008[73]

TABLE 9 Infection and reinfection rates per 100 women-years

		Infection			Re-infection		
Setting	Age (years)	*r*	Rate	*n*[a]	*r*	Rate	*n*[a]
GP	16–17	4	11.2	73	5	86.2	14
	18–20	3	3.1	195	7	22.8	65
	21–24	4	4.3	188	10	26.9	79
FP	16–17	9	9.5	194	13	29.4	95
	18–20	5	3.7	273	12	19.9	127
	21–24	7	7.1	201	5	16.6	63
STD	16–17	5	10.1	102	6	32.3	40
	18–20	16	14.1	235	15	22.8	139
	21–24	9	7.5	245	5	12.8	39.1

a Note that *n* is estimated as the total number of 6-month follow-up periods (events were assumed to happen half way between observations when the rates were estimated in the LaMontagne study[192]). This has been calculated from the reported rates and numbers of events.
Numerators *r* and denominators *n*.
Adapted from tables 2 and 4, LaMontagne *et al.*[192]

Given that observations are interval censored in this study, and the fact that CT infections can clear spontaneously, it is possible for women to have both acquired *and cleared* the infection during the periods between observations, leading to underestimation of incidence. This was not accounted for in the original paper.

Chlamydia trachomatis prevalence by age and setting

Chlamydia trachomatis prevalence information was derived from a recent synthesis of UK data.[106] CT prevalence varies by age and setting. *Table 10* shows estimates of CT prevalence by age in the general population from a logistic regression of UK prevalence studies[106] identified by a systematic review in 2004. Other prevalence data have been collected subsequently[19,81] but have not been incorporated, as it was considered preferable to use results from a single comprehensive study.

Table 11 shows prevalence ORs for the different settings: FP, STD clinics, and general population settings, relative to the GP, from the same study. These are used to inform setting-specific relative risks (RR_s). The interpretation of ORs as RRs is an approximation that is justified by the rarity of the infection.[149]

Statistical models

Model for the overall duration of CT infection in women

The mean duration of infection, Δ, can be expressed as a weighted average of the length of asymptomatic (untreated) infection Δ_A and symptomatic (treated) infection Δ_S:

$$\Delta = \Delta_s \varphi + \Delta_A.(1-\varphi) \tag{10}$$

with φ being the proportion of incident infections in which symptoms develop.

TABLE 10 Estimated prevalence of CT in females in the general population

Age (years)	Prevalence (95% CI)
18–19	0.048 (0.032 to 0.076)
20–24	0.032 (0.021 to 0.049)
25–29	0.015 (0.010 to 0.025)
30–44	0.008 (0.005 to 0.013)

Adapted from table 4, Adams *et al.*[106]

TABLE 11 Adjusted ORs for the effect of setting on chlamydia prevalence in females in the UK

Setting	OR (95% CI)
General population vs. GP	0.6 (0.37 to 0.95)
FP vs. GP	1.27 (1.00 to 1.62)
STD vs. GP	2.39 (0.72 to 3.33)

Adapted from table 3, Adams *et al.*[106]

Regression estimates of infection and re-infection rates by age and setting

We model the infection rates as a function of a baseline infection rate $\lambda_{1,1,1}$ multiplied by the between-setting HRs ρ_s and the between-age group HRs γ_a. The age- and setting-specific re-infection rates $\lambda_{a,s,2}$ equal the respective infection rate multiplied by a setting-specific re-infection HR η_s (equation 11). Other regression models are considered in *Appendix 3*:

$$\lambda_{a,s,1} = \gamma_a \rho_s \lambda_{1,1,1} \text{ and } \lambda_{a,s,2} = \eta_s \lambda_{a,s,1} \tag{11}$$

The infection rates $\lambda_{a,s,i}$ in equation 11 are informed by the data in *Table 9*, which shows the number of initially uninfected women in each age within each setting who were found to be infected after a 6-month follow-up period. However, the mathematical relationship between the infection rates in each group and the proportions infected $\kappa(t)_{a,s,i}$ at the end of a period of time length t is complex. The formula shown in the DAG (*Figure 8*) allows for the fact that in the LaMontagne *et al.* data[192] it is possible for a woman to clear infection spontaneously or through treatment, and then re-acquire infection within the 6-month follow-up. It is necessary, therefore, to take account of the clearance rates of symptomatic and asymptomatic infection, and the proportion of incident infections that become symptomatic (see *Appendix 4*).

Estimation of force of infection

The infection and re-infection rates can be used to estimate the mean force of infection (FOI) $\tilde{\lambda}^{FOI}_{a,s}$ in the CT– women in each setting and age group using equation 12:

$$\tilde{\lambda}^{FOI}_{a,s} = (1 - p_{a,s}).\lambda_{a,s,1} + p_{a,s}.\lambda_{a,s,2} \tag{12}$$

where the weights $p_{a,s}$ are given by the prevalence of CT in each setting observed in the LaMontagne study[192] (see *Appendix 4*). However, as the LaMontagne study[192] samples from only GP, STD and FP settings, it is necessary to turn to a third source of evidence, CT prevalence, to map these estimates of FOI to estimates for the general population.

The CT prevalence data inform the absolute prevalence in 18- to 19-years-olds $\pi_{1,pop}$ (the youngest age group in the study), and the relative risks RR_a of infection in the generic age group a relative to age 18–19 years so that:

$$\pi_{a,pop} = \pi_{1,pop}.RR_a \tag{13}$$

In order to use these data to map our estimates of FOI to the general population, we make the assumption that the between-setting and between-age group RRs in the prevalence data directly inform the HRs (γ_a and ρ_s) in the incidence model described above. For a fixed duration, prevalence ratios must be equal to incidence rate ratios, so the assumption is that ratios of incidence are equivalent to ratios of FOI, and of infection. We consider this assumption further in *Appendix 3*.

A first estimate of *Chlamydia trachomatis* incidence in England (method A)

We use the OR between the general population setting and the GP setting, which informs the parameter ρ_{pop} to map the FOI in the GP setting to provide an estimate of the FOI in women in the general population $\tilde{\lambda}^{FOI}_{a,pop}$:

$$\tilde{\lambda}^{FOI}_{a,pop} = \rho_{pop} \tilde{\lambda}^{FOI}_{a,GP} \tag{14}$$

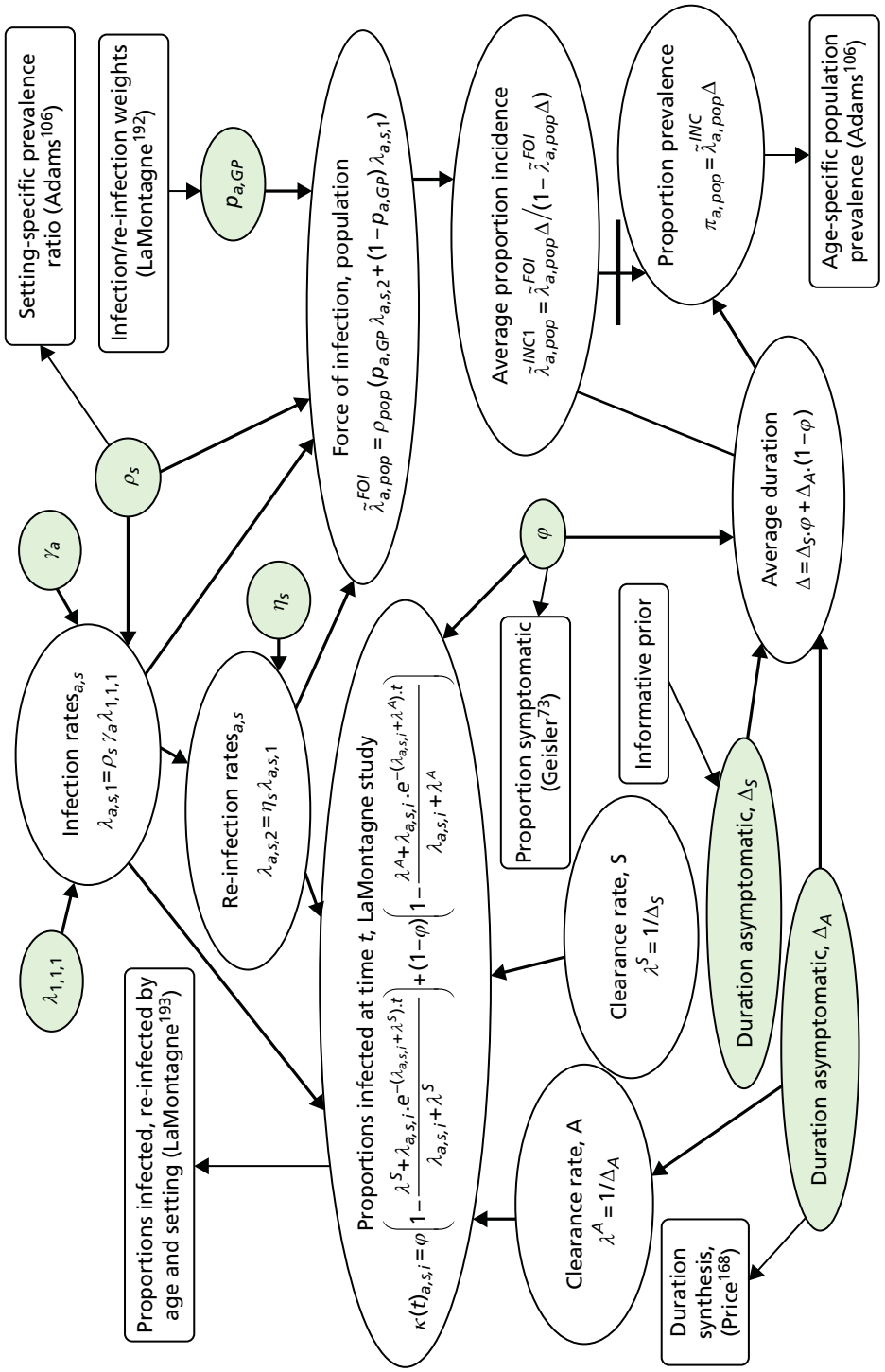

FIGURE 8 Directed acyclic graph of the evidence network.[193] Reproduced from: Price M, Ades A, De Angelis D, Welton NJ, Macleod J, Turner K, *et al.* Incidence of *Chlamydia trachomatis* infection in England: two methods of estimation. *Epidemiol Infect* 2014;**142**:562–7. This work is under the Commons Attribution-NonCommercial license 3.0: http://creativecommons.org/licenses/by-nc/3.0/. The basic parameters are: $\lambda_{1,1,1}$, the infection rate in age group 1, setting 1; γ_a, the HR for infection in age group *a* relative to group 1 (age 16–17 years); ρ_s, the HR for infection in setting *s* relative to setting 1 (GP setting); η_s, the setting-specific re-infection–infection rate ratio; $p_{a,GP}$, the proportion of patients at recruitment in the GP attenders in age group *a* in the LaMontagne study[192] that were in the re-infection group reweighted to account for differential recruitment; Δ_A and Δ_S, the durations of asymptomatic and symptomatic infection; φ the proportion of incident infections in which symptoms develop. The black bar indicates where the network can be cut to obtain separate estimates of incidence (see text).

Estimates of FOI are of interest in themselves. However, we can easily calculate the annual population incidence rate $\tilde{\lambda}^{INC1}_{a,pop}$: (years^{-1}) for the age groups 16–17, 18–20 and 21–24 years as a function of FOI (years^{-1}) and duration using equation 15:

$$\tilde{\lambda}^{INC1}_{a,pop} = \frac{\tilde{\lambda}^{FOI}_{a,pop}}{1-\tilde{\lambda}^{FOI}_{a,pop}\Delta} \tag{15}$$

A second estimate of *Chlamydia trachomatis* incidence in England (method B)

A second estimate of the annual population incidence can be obtained using data on duration and data on prevalence using the relationship incidence = prevalence/duration so that:

$$\tilde{\lambda}^{INC2}_{a,pop} = \frac{\pi_{a,pop}}{\Delta} \tag{16}$$

Where duration is estimated as previously described and prevalence $\pi_{a,pop}$ is informed directly by the data in *Table 10* so $\tilde{\lambda}^{INC2}_{a,pop}$ is estimated for the age groups 18–19, 20–24, 25–29 and 30–44 years.

Full synthesis model

The method A and method B analyses can be combined in a single joint synthesis using the relationship:

$$\pi_{a,pop} = \tilde{\lambda}^{INC}_{a,pop}.\Delta \tag{17}$$

Where $\tilde{\lambda}^{INC}_{a,pop}$ is informed as described in method A, and the parameters $\pi_{a,pop}$ and Δ are informed as described in method B. This is shown in the DAG in *Figure 8*. This single joint analysis provides estimates of population incidence for the age groups 16–17, 18–20, 21–24, 25–29 and 30–44 years. The only age groups for which incidence is estimated in both methods A and B are 18–20 and 21–24 years. However, estimates for the other age groups are also expected to change. The full synthesis model provides estimates for all parameters based on the entire data ensemble. So, for example, when our knowledge of the regression parameters described in method A are updated by the data described in method B, estimates of the annual population incidence rate in 16- to 17-year-olds may change.

Note that the DAG in *Figure 8* also describes methods A and B above. We remove the constraint that *prevalence* = *incidence* × *duration* shown on the DAG under the heavy black bar replacing it with equation 17 above, and place uninformative priors on $\pi_{a,pop}$. This single unconstrained model then produces estimates from both methods A and B in parallel.

Model critique

We compare the goodness of fit of the combined synthesis model and the model with separate incidence estimates: this provides a global assessment of the statistical assumptions. A graphical comparison of the separate incidence estimates is also presented. We assess the validity of some more specific statistical assumptions in *Appendices 3* and *4*. Unless otherwise stated vague priors are employed throughout, so that results are dominated by the data. The WinBUGS code is available along with the data sets in *Appendix 5*. It has been annotated to help readers to understand the model.

Results

Table 12 shows the posterior estimates of the annual population incidence from method A (column 2), method B (column 3) and the full synthesis model (column 4). Estimates from method A are available for the 16–17, 18–19, and 20–24 years age groups, method B provides estimates for age groups 18–19, 20–24, 25–29 and 30–44 years, and estimates for all age groups are available from the full synthesis model. Estimates from the full synthesis model are around a factor of 1.5 lower than those obtained from method A but only marginally higher than those obtained from method B. This is because the uncertainty in the incidence information from the LaMontagne data[192] is much greater than in the combined duration and prevalence information. This effect is shown pictorially in *Figure 9*, which compares the estimates of incidence in the age groups 18–19 years and 20–24 years, and also shows the combined estimate incorporating all data sources. Results from the full synthesis model for all five age groups are also given in *Figure 10*.

Table 13 shows the estimates of the basic parameters in the model estimated when the constraint that *prevalence* = *incidence* × *duration* is excluded (column 2) and included (column 3) in the model, representing, respectively, methods A and B being performed simultaneously in parallel, and the full synthesis model. It shows that for most parameters: duration, proportion symptomatic, re-infection-to-infection rate ratios, age- and setting-specific risk ratios, and prevalence parameters, the synthesis has not contributed much additional information over and above the 'direct' data already available. However, the general population to GP RR is lowered by a factor of about 1.35 compared with method A, and the 95% CrIs are about half of the width.

Our separate models of incidence in women aged 16–24 years included nine parameters and had a residual deviance of 18.4 for a data set with 21 data points (see *Tables 9–11*) representing a good fit (*Table 14*). When these data are combined with the prevalence information, residual deviance increases only marginally (19.5), indicating a lack of conflict between the different sources of information on incidence. Prevalence and duration data also fitted equally well. Results (not shown) with a wider uniform prior distribution on the duration of symptomatic infection, 3–12 weeks rather than 4–8 weeks, were almost identical (< 1% multiplicative change). We therefore recommend using results from the full synthesis model that uses all of the data, giving an estimated incidence rate in 16- to 24-year-old females, the population targeted by the NCSP, of 0.05 per year (95% CrI 0.035 to 0.071) and in 16- to 44-year-olds 0.021 per year (95% CrI 0.015 to 0.028).

TABLE 12 Population CT incidence rate per year in women by age estimated using each method. From Price *et al.*[193]

Parameter (years)	Method A (adjusted incidence data) (95% CrI)	Method B (prevalence and duration data) (95% CrI)	Full synthesis model (95% CrI)
$\tilde{\lambda}^{INC}_{16-17,pop}$ incidence, 16–17	0.122 (0.057 to 0.235)	–	0.082 (0.047 to 0.134)
$\tilde{\lambda}^{INC}_{18-19,pop}$ incidence, 18–19	0.070 (0.036 to 0.126)	0.046 (0.028 to 0.072)	0.048 (0.032 to 0.068)
$\tilde{\lambda}^{INC}_{20-24,pop}$ incidence, 20–24	0.060 (0.032 to 0.106)	0.031 (0.019 to 0.048)	0.039 (0.027 to 0.054)
$\tilde{\lambda}^{INC}_{25-29,pop}$ incidence, 25–29	–	0.015 (0.009 to 0.023)	0.015 (0.0087 to 0.024)
$\tilde{\lambda}^{INC}_{30-44,pop}$ incidence, 30–44	–	0.0078 (0.0045 to 0.013)	0.0080 (0.0045 to 0.013)
Incidence, 16–24			0.050 (0.035 to 0.071)
Incidence, 16–44			0.021 (0.015 to 0.028)

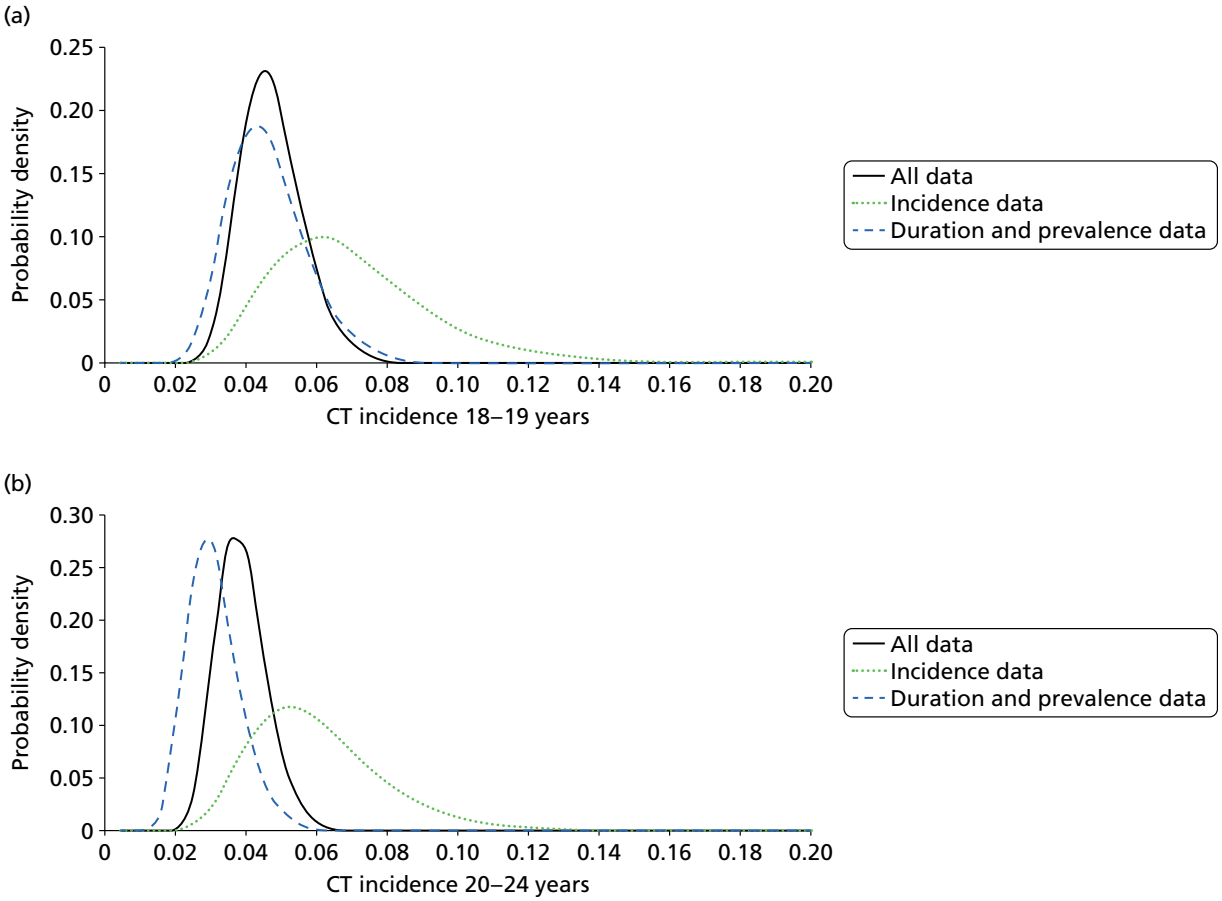

FIGURE 9 Posterior distributions of incidence parameters, comparing results based on incidence data, with results based on information in prevalence and duration studies; (a) age group 18–19 years, (b) age group 20–24 years.[193] Reproduced from Price M, Ades A, De Angelis D, Welton NJ, Macleod J, Turner K, *et al.* Incidence of *Chlamydia trachomatis* infection in England: two methods of estimation. *Epidemiol Infect* 2014;**142**:562–67. This work is under the Commons Attribution-NonCommercial license 3.0: http://creativecommons.org/licenses/by-nc/3.0/.

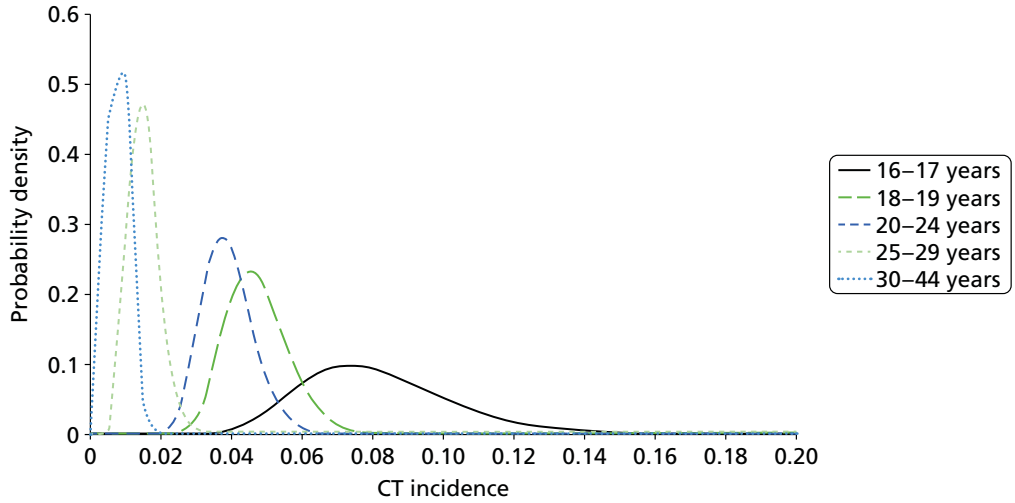

FIGURE 10 Posterior distribution of incidence, by age range, based on all available information.

TABLE 13 Incidence, prevalence, duration estimates obtained from the separate models, and from the full synthesis model. From Price *et al.*[193]

Parameter	Separate models:[a] mean (95% CrI)	Full synthesis model: mean (95% CrI)
CT duration and clearance rate		
λ^A clearance rate, asymptomatic	0.74 (0.62 to 0.90)	0.77 (0.64 to 0.95)
Δ_A duration of asymptomatic infection	1.36 (1.13 to 1.63)	1.31 (1.06 to 1.56)
Δ mean duration (years)	1.07 (0.86 to 1.29)	1.03 (0.82 to 1.25)
Φ proportion symptomatic	0.23 (0.16 to 0.31)	0.23 (0.16 to 0.32)
CT incidence: regression parameters		
η_{GP} re-infection: infection ratio, GP	7.31 (4.07 to 11.9)	7.08 (3.97 to 11.6)
η_{FP} re-infection: infection ratio, FP	3.52 (2.09 to 5.52)	3.66 (2.16 to 5.77)
η_{STD} re-infection: infection ratio, STD	2.01 (1.17 to 3.17)	2.08 (1.21 to 3.28)
ρ_{FP} HR, GP (reference group)	1	1
ρ_{pop} HR, general population	0.62 (0.37 to 0.96)	0.46 (0.33 to 0.63)
ρ_{FP} HR, FP	1.28 (1.02 to 1.59)	1.30 (1.03 to 1.61)
ρ_{STD} HR, STD	2.38 (1.78 to 3.11)	2.45 (1.83 to 3.20)
CT prevalence, %		
$\pi_{16–17,pop}$ general population, 16–17 years	–	8.38 (4.94 to 13.5)
$\pi_{18–19,pop}$ general population, 18–19 years	4.91 (3.13 to7.28)	4.85 (3.47 to 6.59)
$\pi_{20–24,pop}$ general population, 20–24 years	3.27 (2.10 to 4.83)	3.96 (2.89 to 5.30)
$\pi_{25–29,pop}$ general population, 25–29 years	1.54 (0.95 to 2.35)	1.54 (0.95 to 2.37)
$\pi_{30–44,pop}$ general population, 30–44 years	0.82 (0.50 to 1.28)	0.83 (0.50 to 1.29)
General population, 16–24 years		5.15 (3.76 to 6.93)
General population, 16–44 years		2.11 (1.63 to 2.71)

a Methods A and B.

TABLE 14 Model fit statistics for each data set from the separate models fitted using methods A and B in parallel and the full synthesis model. From Price *et al.*[193]

		Number of parameters		Mean residual deviance	
Source	Number of data points	Separate models	Full model	Separate models	Full model
Incidence	21	9	9	18.4	19.5
Prevalence	4	4	2	4.0	4.1
Duration	2	2	2	2.0	2.2
Total	27	15	13	24.3	25.7

Discussion

Although CT prevalence in the general UK population has been studied,[106] incidence estimates have been produced only in clinic patients.[192] We used data on ratios between clinic settings and the general population in prevalence studies to 'recalibrate' the incidence data to a lower value that was appropriate to a general population setting. The possibility of clearance of infection and re-infection during the follow-up period was also taken into account: the effect of this is to raise incidence estimates above the levels that are directly observed. The estimate based on the recalibrated incidence study was found to be compatible with an estimate based on combining prevalence and duration information.

A certain degree of simplification is involved. The incidence data were collected in 2003–4, 2–3 years later than the NATSAL (National Survey of Sexual Attitudes and Lifestyles) study,[195] which contributes all of the general population prevalence information to the estimates in the Adams study,[106] and in a period before intensive screening was taking place. We have assumed that incidence is unlikely to have changed greatly between these dates, and that our estimates are therefore relevant to the years 2001–5.

As with any evidence synthesis method, conclusions are limited by the quality of the original data and the assumptions made in interpreting them. The CT prevalence information in NATSAL[195] was based on self-testing in a structured population sample and is vulnerable to response biases, although these have been extensively analysed elsewhere.[196] The incidence data were collected in two English areas, which were metropolitan and urban, respectively. The extent to which these data can be assumed to be nationally representative is not known. The final estimates are more influenced by the prevalence data than the incidence data, and this is dominated by the NATSAL[195] data collected in 2000. Finally, the estimates of duration of asymptomatic CT duration, 1.36 (95% CrI 1.11 to 1.62) years, are based on the *Chapter 4* synthesis of studies with different designs and subject to the assumptions made in that chapter.[168]

Additional validation of our estimates and of our overall approach is available by multiplying our estimated infection rate by the number of women aged 16–24 years in England, based on population census projections for 2002.[197] This predicts a total of 137,100 (95% CrI 95,520 to 192,500) infections. This can be compared with the 31,510 and 34,660 women aged 16–24 years who were treated for CT in STD clinics in 2002 and 2003, respectively.[198] The ratios of numbers treated to predicted total infections in women aged 16–24 years are 24% (95% CrI 16% to 33%) for 2002 and 26% (95% CrI 18% to 36%) for 2003. This accords closely with the proportion of infections in which symptoms develop estimated from the model, and the Geisler *et al.*[73] findings. Therefore, had we partitioned women as treated or untreated when estimating the mean duration as is often done in dynamic models and used recursive equations to estimate the proportion treated from routine data, we would have obtained almost identical results.

Although the estimates of CT incidence and prevalence in England may be of limited interest elsewhere, the study does have wider implications. First, the fact that incidence, prevalence and duration evidence is internally consistent provides a degree of independent validation of our estimates of all three parameters. Second, it indicates that estimates of CT prevalence *or* CT incidence in other countries can each be generated from the other, using our estimates of duration. Alternatively, where information is available on both incidence and prevalence, a similar exercise could be carried out to provide a further validation of our results and the models on which they are based.

The study raises the question: what is the best way to obtain accurate population-based estimates of CT incidence? Further direct study of infection and re-infection rates among opportunistically recruited women appears to be worthwhile. However, as well as taking account of clearance and re-infection during follow-up, it will probably always be necessary to 'recalibrate' setting-specific estimates to the general population. Studies of either prevalence or incidence based on structured general population surveys are, therefore, essential.

Our analysis of incidence, prevalence and duration has relied on an essentially static epidemiological model. The alternative would be to assess the consistency of a somewhat wider evidence base within a dynamic modelling context. For example, a dynamic model could be estimated from the same sources of data (incidence, prevalence, duration of symptomatic and asymptomatic infection, proportion symptomatic), but also incorporating information on contact rates and transmission rates per contact (see *Chapter 12*).

Summary of assumptions and findings

Key assumptions

1. The LaMontagne study[192] of incidence in FP, STI and GP groups is nationally representative.
2. The incidence and prevalence estimates are based on several data sources, but are relevant to the years 2000–5.

Summary of findings

1. The available data on incidence in the UK agree with predictions based on duration (see *Chapter 4*) and UK prevalence.
2. The mean duration of CT infection in women (including symptomatic and asymptomatic infections) is 1.03 years (95% CrI 0.82 to 1.25 years). The proportion of incident infections that are symptomatic is 0.23 (95% CrI 0.16 to 0.32).
3. CT incidence in the general population aged 16–24 years is 5.04 per 100 person-years (95% CrI 3.52 to 7.09) and in women aged 16–44 years it is 2.07 per 100 person-years (95% CrI 1.50 to 2.81).
4. CT prevalence in the general population aged 16–24 years is 5.15% (95% CrI 3.76% to 6.93%), and in women aged 16–44 years it is 2.11% (95% CrI 1.63% to 2.71%).

Chapter 6 Risk of pelvic inflammatory disease following *Chlamydia trachomatis*: analysis of prospective studies

Objectives

To:

1. develop and apply appropriate models for trials of screening and other prospective studies that follow-up women with CT to observe whether they develop PID
2. estimate the risk of PID following a CT infection
3. estimate the proportion of CT-related PID infection in women with CT that can be prevented by CT screening and treatment at annual intervals.

Introduction

A range of observational studies have followed untreated CT infected (CT+) women prospectively and observed the number of women developing PID.[73,165,184,199,200] In some studies,[73,184,199,200] the infection, presumably mostly symptomatic, has been diagnosed in a clinic, and in one study[165] it was identified by screening asymptomatic cases. There have also been controlled trials randomising large numbers of women to screening and treatment or to no screening.[12,13,19] However, as we note earlier, there is a remarkable lack of consensus on the risk of PID attributable to an episode of CT, despite a major authoritative review.[98]

Previous reviews[20,61,98,144] of prospective data on progression to PID have been qualitative. In this chapter we synthesise data from these studies to provide quantitative estimates of the difference in the rates at which PID develops in women with and without CT. From these estimates we derive first the probability that an episode of CT causes an episode of PID, and then the proportion of episodes of PID caused by CT in women with CT, that could be prevented by screening at 1-yearly intervals. By this we mean annual screening not associated with recent risk of infection.

First, we re-appraise the prospective studies, including non-randomised and uncontrolled studies. We then present a three-state continuous time Markov model, which allows not only for different rates of development of PID in CT-infected (CT+) and uninfected (CT–) women but also spontaneous clearance of CT, a competing risk. It also allows for CT infection in women who are initially *un*infected (CT–), and for *re*-infection with CT in women who are initially CT+ but who spontaneously clear infection or are tested and treated. We then re-analyse each of the controlled studies, both randomised and not, under the assumptions of the model. Finally, we synthesise them altogether to obtain estimates of the proportion of incident CT that results in PID, and the proportion of CT-caused PID in women with CT that can be prevented by annual screening. The work in this chapter has been published previously.[201]

The definition of PID that was used in the prospective studies is, of course, a key issue in their interpretation. Our assumption is (see *Chapter 2*) that, in prospective studies, the 89% of PID that is not asymptomatic[103] will be picked up.

Methods

Evidence identification

We sought prospective studies of women with CT including observational studies with or without a control group, and RCTs of screening interventions. We identified studies from recent review papers,[20,61,98,144] reference searches of published economic and population transmission models,[1,2,8,29,140,141,170–174] and as part of a wider review of the literature on the natural history of CT. Included among these was a recent systematic review instituted at an Expert Advisory Meeting of the Centers for Disease Control and Prevention.[98] We exclude the study by Bachmann *et al.*[202] because the follow-up time was not reported; the study by Stamm *et al.*[203] because the patients were co-infected with gonorrhoea; and the study by Westergaard *et al.*[204] because the patients were women who had just undergone an abortion, and this is a very unrepresentative population. The study by Morre *et al.*[180] is excluded because of uncertainty about the adequacy of outcome assessment.[40] We identified eight studies[12,13,19,73,165,184,199,200] (*Table 15*) that followed up women with CT to assess the proportion who developed PID in the absence of, or prior to, treatment. Note that four of the studies[73,165,184,200] shown in *Table 15* are excluded from the modelling because of a lack of a control group.

Information sources and their interpretation

Some studies were set in STD clinics, and others were based on infected women recruited by population screening. Five were observational, one of which, a clinic-based study, included a control group. The remainder were RCTs, all based on screened populations. In the POPI trial,[19] both arms were screened; one was treated immediately, whereas in the other arm treatment was deferred until the end of the study.

Table 15 shows the proportion of women with PID in each arm at the end of each study. Most investigators have focused attention on these proportions, and in controlled studies, on RRs. However, it is necessary to account for the mean follow-up time t_o for each observation o. The last column of *Table 15* shows crude study-specific estimates of the rates of progression to PID, based on a constant hazard within each study and arm.

The crude estimates of the progression rate in CT+ women vary by factors of 30-fold: studies based in STI clinics having shorter follow-up periods and higher apparent progression rates, whereas those based on screened populations tend to have longer follow-up periods and lower progression rates. There are three possible contributory explanations for this. First, CT infections spontaneously resolve without treatment, so even if progression rates were constant over time, with CT clearance as a 'competing risk', the numbers acquiring PID must slowly decrease over time towards a background level seen in a representative control group. The effect of this will tend to be greater in studies with a longer follow-up. However, the mean duration of untreated CT has been estimated to be over 1 year[74,142,166] (see *Chapter 4*) so this would explain only a small part of the apparent differences. A second possible explanation is that women visiting STI clinics are more likely to have symptoms than those identified through screening, and the higher CT load associated with symptoms[205–207] may confer a greater risk of developing PID. A third possible explanation is if the risk of developing PID is greater in the period immediately after CT infection,[98,208] even if the infection is asymptomatic.

There is a fundamental difference between STI clinic studies and screening studies. Women are most likely to attend clinics because of the onset of symptoms, or because of concern following recent possible exposure to infection. They are therefore more likely to be recently infected. This is supported by evidence that symptoms mostly occur soon after infection (6) (see *Chapter 4*). Screening studies, by contrast, recruit participants with asymptomatic, prevalent CT infection, who have already 'survived' an unknown period of time without developing PID. Further, if the risk of PID is higher immediately following infection, the screening studies must underestimate the overall progression rate. For example, if asymptomatic infection lasted exactly 12 months, and if 20% of incident cases develop PID (at which point they are treated for CT), always exactly 3 months after infection, the annual progression rate that would be observed in patients recruited from a prevalent population would be only 5.9% [i.e. $3/(0.8 \times 12 + 0.2 \times 3) \times 0.2$],

TABLE 15 Information from prospective studies that followed women with chlamydia to PID. From Price *et al.*[201]

First author	Study design	Study population	Arm	Data: r_o/n_o	Crude proportion developing PID (95% CrI): $p_o = r_o/n_o$	Follow-up period: t_o	Crude annual PID rate (95% CrI): $(-\log(1-p_o))/t_o$
Hook[200]	Uncontrolled	Clinic	CT+	3/93	0.032 (0.012 to 0.091)	2 weeks	0.82 (0.30 to 2.40)
Geisler[73]	Uncontrolled	Clinic	CT+	2/115	0.017 (0.005 to 0.061)	2 weeks	0.44 (0.14 to 1.59)
Paavonen[184]	Uncontrolled	Clinic	CT+	3/15	0.200 (0.078 to 0.481)	1 month	2.79 (1.01 to 8.21)
Rahm[165]	Uncontrolled	Screened	CT+	4/102	0.048 (0.016 to 0.097)	3 months	0.18 (0.07 to 0.46)
Rees[199]	Controlled	Clinic	CT+	8/67	0.119 (0.063 to 0.222)	7–90 days	1.02 (0.52 to 2.01)
			CT–	3/62	0.048 (0.018 to 0.135)		0.40 (0.15 to 1.16)
Oakshott[19]	RCT	Screened	Untreated CT+	7/74	0.093 (0.047 to 0.183)	1 year	0.099 (0.049 to 0.205)
			Delayed screen CT–	16/1112	0.014 (0.009 to 0.023)		0.014 (0.009 to 0.024)
			Treated CT+	1/63	0.016 (0.004 to 0.087)		0.016 (0.004 to 0.089)
			Screened CT–	14/1128	0.011 (0.007 to 0.019)		0.012 (0.007 to 0.021)
Scholes[13]	RCT	Screened	Unscreened	33/1598	0.021 (0.015 to 0.029)	1 year	0.021 (0.015 to 0.029)
			Screened	7/645	0.011 (0.005 to 0.022)		0.011 (0.005 to 0.022)
			Chlamydia prevalence	44/645	0.068 (0.051 to 0.091)		
Ostergaard[12]	RCT	Screened	Unscreened	20/487	0.041 (0.027 to 0.063)	1 year	0.042 (0.027 to 0.065)
			Screened	9/443	0.020 (0.011 to 0.038)		0.021 (0.011 to 0.039)
			Chlamydia prevalence	43/867	0.050 (0.037 to 0.066)		

r_o, n_o, t_o are the numerator, denominator and time interval in observation o.

a gross underestimate. In such a scenario, only screening and treatment occurring during the first 3 months of a 12-month infection would prevent PID due to these infections. As a result, a single screen every year would prevent only 29.4% (5.9/20% = 29.4%) of the PID cases caused by CT.

A model for progression to pelvic inflammatory disease in the controlled prospective studies

In our model for prospective studies (*Figure 11*), women who are CT+ and CT– at the start of the study begin in states 1 and 2, respectively. Women in state 1 may clear infection and progress to the CT– state, state 2, at rate λ^C, or they may develop PID (state 3) at a study-specific rate θ_s^{CT+} for study *s*. Women in state 2 may become infected with CT, which occurs at rate λ_s^I or may develop PID at rate θ_s^{CT-} if they acquire, or already had, a non-CT infection that carries a PID risk. We assume that the prevalence and incidence rates for other STIs are the same in each arm. The difference δ between θ_s^{CT+} and θ_s^{CT-} represents the rate of acquiring PID in women with a current CT infection that can be causally attributed to CT. The *Figure 11* model is represented by the Markov transition rate matrix (G) shown in equation 18:

$$G = \begin{bmatrix} -(\lambda^C + \theta_s^{CT+}) & \lambda^C & \theta_s^{CT+} \\ \lambda_s^I & -(\lambda_s^I + \theta_s^{CT-}) & \theta_s^{CT-} \\ 0 & 0 & 1 \end{bmatrix} \quad (18)$$

Only comparative studies can contribute information on δ but we cannot derive estimates of δ from comparative studies without information on λ^C and λ_s^I. We assume that the CT clearance rate for asymptomatic infection, λ^C, is constant across studies, at 0.74 (95% CrI 0.61 to 0.89) per year, based on the *Chapter 4* evidence synthesis of studies on CT duration. Similar estimates have been reported elsewhere.[166] Note that the information on incidence and prevalence can back-propagate through the evidence network and impact on the posterior distribution of the duration of asymptomatic infection, although the latter was not materially changed.

For CT incidence λ_s^I we assume an infection rate of 5% per year in women who are CT– at the outset, and a re-infection rate of 15% per year in women who are CT+ and tested and treated. Given a CT average duration of 1–1.5 years, a 5% incidence reflects a CT prevalence of 5–7.5%, which accords with the baseline CT prevalence in the three trials. Re-infection rates are higher by approximately a factor of 3.[192] We assume that all women in the Scholes[13] and Ostergaard[12] (95% CT–) studies are at risk of infection, whereas all women in the Rahm[165] (50% CT+) and POPI[19] (100% CT+) studies are assumed to be at risk of re-infection. Detailed sensitivity analyses (see *Appendix 6*) show that conclusions are relatively insensitive to the values assumed.

Some of the screening studies record the proportion of women who were tested and treated for CT during follow-up, thus shortening their time at risk of CT-related PID. This was 43% for the POPI trial,[19] 32% for Ostergaard *et al.*[12] and 8.1% for Rees.[199] The Scholes *et al.* paper[13] implies that some women were tested but does not report how many: we have assumed a proportion of 32%, following Ostergaard *et al.*,[12] but the impact of a lower rate is explored in sensitivity analyses (see *Appendix 6*).

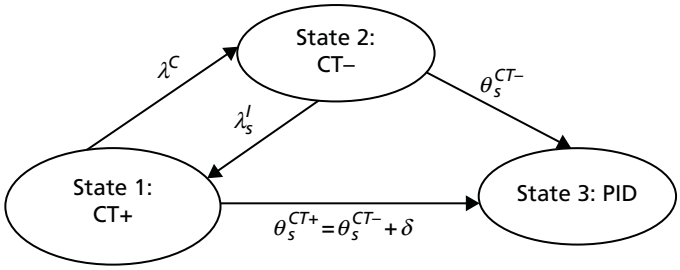

FIGURE 11 Multi-state transition model for disease progression in prospective studies. From Price *et al.*[201]

We consider two models. In the first, the causal rate of PID due to CT (δ) is constant throughout the duration of infection. In the second, the causal rate in the period immediately following CT infection (δ_1) is different from the causal rate during the remainder of the infection (δ_2). We fit models with the rate δ_1 lasting for periods of 30 days, 60 days and 90 days post infection.

Estimation

Details of the estimation methods are given in *Appendix 7*. In brief, each study arm provided data with a binomial likelihood, which estimates a probability parameter, say $p_{s,t}^{i \to j}$, representing the study-specific transition probability that a person in state *i* occupies state *j*, *t* years later. These probabilities are entries in the transition probability matrix P_t. If the transition rates are constant over the period *t* then the transition probabilities, on which we have data (see *Table 15*), are functionally related to the transition rates in equation 18 by Kolmogorov's Forward Equations (equation 19):[209]

$$\frac{dP_t}{dt} = P_t.G \tag{19}$$

Estimation is carried out using a Bayesian approach, where the posterior distribution is sampled through Markov chain Monte Carlo (MCMC) simulation. The rate parameters in equation 18 are given vague priors, the values of which are updated by the data. Numerical solutions to the Forward Equations are found in each MCMC cycle, using the WinBUGS[44] add-on WBDiff,[161] which uses the Runge–Kutta method.[210,211]

Only the CT+ groups from the POPI trial[19] were included in the analysis because the treatment effect estimate is thereby based on a randomised comparison. The proportion with CT in the early treated arm are considered, in effect, as being CT– at the start of the observation period. Those in the deferred treatment group are, of course CT+. However, in the Scholes[13] and Ostergaard[12] trials, although the women in the screened arm are again taken to be CT– at the start of follow-up, those in the unscreened arm are a *mixture* of (untreated) CT+ and CT–. The study-specific CT prevalences in the screened arms are estimators of these mixing proportions.

Further details on how the observed data informs the rate parameters in the matrix G (equation 18) are given in *Appendix 7*. The joint posterior distribution of the rate parameters is used to obtain posterior distributions for the two parameters of interest: the proportion of CT infections that result in PID and the proportion of CT-caused PID preventable by annual screening.

Probability that incident *Chlamydia trachomatis* will cause an episode of pelvic inflammatory disease

Although the comparative studies provide an estimate of the difference δ in progression rates between CT+ and CT– women, the target parameter of interest is the proportion of incident CT infections that cause PID. We assume that a proportion, φ, of incident CT cases are symptomatic and treated, clearing at rate λ^T, and with no increased risk of PID. The remaining $1-\varphi$ infections are asymptomatic. To simplify the equations and maintain generality we ignore non-CT-related PID as a competing risk at this stage. Even if the rate of non-CT PID were twice that observed in the screened arm of POPI,[19] this would change the value of our target parameter by a multiplicative factor of only about 2.5%. The probability that either type of infection will cause PID is the ratio of the rate of leaving state 1 for state 3 to the rate of leaving state 1, a standard result from competing risks analysis.[212] The proportion of incident cases of CT leading to an episode of PID in a *homogeneous model* is therefore:

$$\kappa = (1-\varphi).\frac{\delta}{\delta+\lambda^C} + \varphi.\frac{\delta}{\delta+\lambda^T} \tag{20}$$

Calculation of κ requires estimates of φ and λ^T. For the former, we use the estimate from Geisler *et al.*:[73] 0.24 (95% CrI 0.17 to 0.32). For the latter we assume that treated infection lasts 4–8 weeks.

In the *two-stage piecewise homogeneous model*, with the initial rate lasting for B days, the probability of developing PID in the first B days and the probability of developing PID subsequently conditional on having neither developed PID nor cleared infection in the first B days, are summed (equation 21):

$$\begin{aligned}\kappa = (1-\varphi).\left(1-\exp\left(-(\lambda^C+\delta_1).\frac{B}{365}\right).\left(\frac{\delta_1}{\delta_1+\lambda^C}\right)+\exp\left(-(\lambda^C+\delta_1).\frac{B}{365}\right).\left(\frac{\delta_2}{\delta_2+\lambda^C}\right)\right)\\ +\varphi.\left(1-\exp\left(-(\lambda^T+\delta_1).\frac{B}{365}\right).\left(\frac{\delta_1}{\delta_1+\lambda^T}\right)+\exp\left(-(\lambda^T+\delta_1).\frac{B}{365}\right).\left(\frac{\delta_2}{\delta_2+\lambda^T}\right)\right)\end{aligned} \quad (21)$$

Estimation of the proportion of pelvic inflammatory disease episodes prevented by screening

Even with 100% coverage, an annual screening programme for CT would not prevent all cases of PID. We estimate the proportion of episodes of CT-related PID that could be prevented in women who become infected with CT. The expressions for this quantity are derived and presented in *Appendix 7*. This applies to the benefits to CT+ women who are screened and treated, and is not designed to measure the full effect of screening as it does not take into account reduction in PID due to reduced CT incidence.

Plan of analysis

The data analysis is carried out in several steps in order to show the influence of the different types of data on the estimates of interest. First, we examine each of the comparative studies separately. We then synthesise information from all three RCTs, and then from the four comparative studies. Studies are combined on the basis that the progression rates in CT– women vary between studies in an arbitrary way, but the difference δ between the progression rates from CT– and CT+ has the same fixed effect across studies. WinBUGS code is provided in *Appendix 8*.

Results

Table 16 shows the study-specific estimates from the homogeneous Markov model of the difference δ, which we interpret as the causal rate of PID in CT+ women, and the proportion of incident CT infections that result in PID. Although the POPI trial[19] data delivered the lowest estimates of both parameters, the CrIs for the estimates from the other studies were considerably wider and approximately compatible with the POPI[19] results. The estimates from the Scholes study[13] were relatively insensitive to changes in assumptions on the proportion tested during the follow-up period (see *Appendix 6*).

TABLE 16 Posterior means and 95% CrIs for the causal rate of PID in CT+ women, and the probability that an incident CT case causes PID.[201] Reproduced from Price M, Ades A, De Angelis D, *et al.* Risk of Pelvic Inflammatory Disease following *Chlamydia trachomatis* infection: analysis of prospective studies with a multistate model. *Am J Epidemiol* 2013;**178**:484–92

Study	Causal PID rate in CT+, δ	Pr(CT causes PID)
POPI	0.15 (0.02 to 0.31)	0.12 (0.02 to 0.24)
Scholes	0.26 (0.03 to 0.55)	0.20 (0.03 to 0.34)
Ostergaard	1.03 (0.13 to 2.49)	0.42 (0.11 to 0.64)
Rees	0.73 (0.06 to 1.71)	0.36 (0.06 to 0.58)

Table 17 presents the results of combining the information from the different studies under each of the homogeneous and piecewise homogeneous models. In all of the models the mean residual deviance was close to the number of data points, indicating a good fit. Also, it was apparent that the estimates under the one-rate model were broadly similar for data (1) from the POPI trial[19] only, or (2) pooled from all trials, or (3) including the observational study as well. The causal rate of PID ranged from 0.15 to 0.19 per year, and the probability of CT causing PID from 0.12 to 0.16. The proportion of CT-caused PID preventable by annual screening remained at 61–63%. The degree of overlap between these estimates, and the shape of the posterior distributions can be seen in *Figure 12*.

The data did not distinguish between the one- and two-rate models in terms of goodness of fit. The rate in the initial period is estimated to be approximately 50% higher than the rate during the remainder of infection, which is slightly lower than for the one-rate model. Both rates are insensitive to the assumed time period of the higher rate between 30 and 90 days (see *Table 17*). Credible intervals are, however, very wide, and a Bayesian *p*-value, testing the null hypothesis of equal rates in the 60-day model, was only 0.67. The assumption of two rates had no real effect on the causal probability of PID from an episode of CT. The estimated proportion of PIDs that can be prevented by screening reduced marginally to between 54% and 58% for B = 90 days and 30 days, respectively, although CrIs were wide (*Figure 13*).

Sensitivity analyses (see *Appendix 6*) showed that although results from some of the individual studies can be sensitive to changes in assumed infection and re-infection rates, none of the results for the pooled analyses using the single rate model was changed by more than a multiplicative factor of 4%. Results from the two-rate model were more sensitive, with κ varying between 0.15 and 0.20 when B = 60 days.

TABLE 17 Synthesis models. Residual mean deviance, numbers of data points and marginal posterior means with 95% CrIs for key parameters. From Price *et al.*[201]

Method	Mean residual deviance	Number of data points	Causal rate of PID	Probability CT causes clinical PID	Proportion prevented by screening
One-rate models					
POPI data only	2.0	2	δ 0.145 (0.021 to 0.31)	0.125 (0.021 to 0.24)	0.625 (0.56 to 0.69)
Trials only	8.5	8	δ 0.178 (0.049 to 0.33)	0.148 (0.048 to 0.25)	0.615 (0.55 to 0.68)
All controlled studies	10.6	10	δ 0.189 (0.057 to 0.34)	0.156 (0.056 to 0.25)	0.612 (0.55 to 0.67)
Two-rate models					
All controlled studies, 30 days	10.3	10	δ_1 0.280 (0.016 to 0.65) δ_2 0.187 (0.055 to 0.35)	0.161 (0.068 to 0.26)	0.576 (0.42 to 0.68)
All controlled studies, 60 days	10.3	10	δ_1 0.288 (0.018 to 0.65) δ_2 0.182 (0.046 to 0.35)	0.163 (0.073 to 0.26)	0.552 (0.32 to 0.72)
All controlled studies, 90 days	10.3	10	δ_1 0.281 (0.017 to 0.63) δ_2 0.179 (0.038 to 0.36)	0.165 (0.075 to 0.26)	0.543 (0.26 to 0.76)

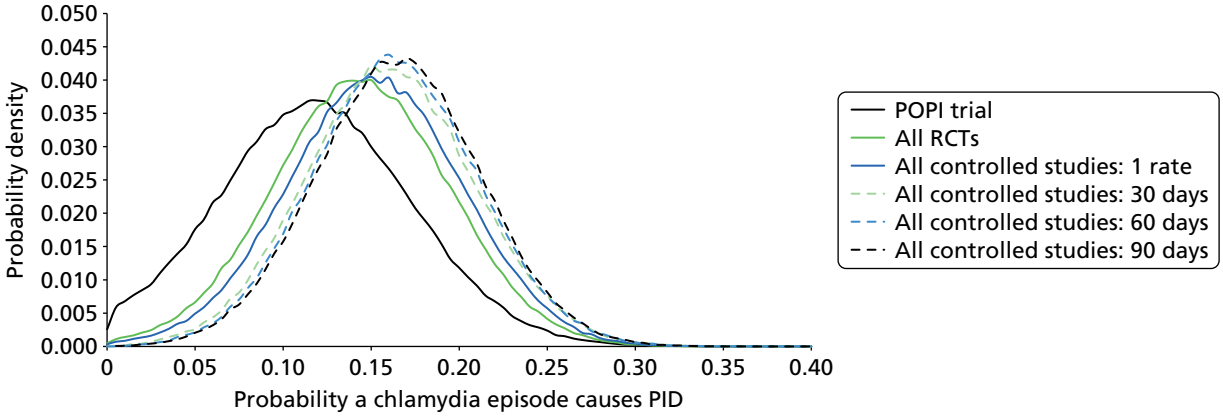

FIGURE 12 Marginal posterior distributions of the probability a CT infection causes clinical PID, estimated by each model. From Price *et al.*[201]

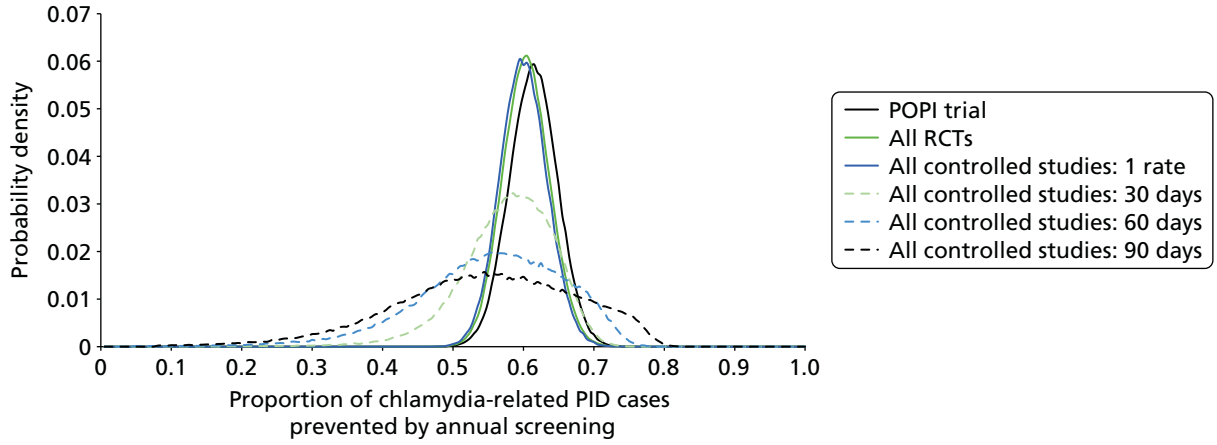

FIGURE 13 Marginal posterior distributions of the proportion of CT-related PID episodes, in women with CT, prevented by annual screening. From Price *et al.*[201]

Discussion

Early literature on the progression from CT to PID cited high rates based on uncontrolled studies in high risk populations.[184,203] Over time, these estimates were increasingly questioned.[5,39,213] Quantitative estimates of the causal relation between CT and PID have been made from a wide range of study types,[40] including linkage studies of treated women,[36,107,214–217] arguments from retrospective data[1,36,107,214–217] and the use of CT serology.[88,101,218] More recently, it has been accepted that linkage studies cannot be used to generate reliable estimates.[20,98]

The approach presented here has some novel aspects. The Markov model characterises the causal rate as the difference in progression rates in CT– and CT+ individuals. It takes account of the 'competing' role of CT clearance and of the fact that persons testing CT+ upon screening have already 'survived' a period without developing PID. Finally, we have allowed for the impact of CT treatment of women in the untreated arms of trials, and probable re-infection rates among those who were treated.

The continuous time Markov formulation avoids the limitations of traditional RR estimates[12,13,19] of the effect of CT screening. As the follow-up period is extended, the risk ratio estimate will be diluted by CT clearance and PID caused by infections (both CT and non-CT) acquired after the screening and treatment, while in time the risk difference should converge to a stable estimate. Crude risk difference estimates[19,219] should also be interpreted with caution. The POPI trial[19] data suggest that screening can prevent 83% of

PID occurring in the following year in a *prevalent* CT+ population (11). Based on the Markov model, these data tell us that the proportion of *incident* CT-caused PID that can be prevented by annual screening is lower, between 55% and 65%.

The study has also shown for the first time that similar estimates can be derived, though indirectly and with less precision, from trials of screening versus no screening. The Scholes trial in particular has been criticised[18] because testing and treatment rates outside the protocol are unlikely to be the same in the two arms. We have used the information available, and sensitivity analyses, to mitigate this. Even with this defect, the trials are clearly superior sources of evidence on the causal effects of CT to the often-cited observational studies, in which the difference in CT status at the outset is likely to correlate with risk factors that favour development of PID: higher partner change, greater risk of exposure to other agents that cause PID, and greater prior exposure to CT.

The study relates to 'annual screening' at a random point in time. In practice, the benefit of screening for the individual infected women may be greater if it occurs soon after partner change, as recommended in some countries, or less if partner testing and treatment are not pursued effectively.

The main weakness of this analysis lies in the reliance of almost all prospective studies on clinical PID as the outcome and between-study variation in case definition.[40] There is a clear risk that prospective studies may *under*diagnose PID, as case–control studies of TFI suggest that up to 80% of women with TFI report no history of PID.[34,99,220] However, Wolner-Hanssen[103] found that only 11% of women with TFI reported *never* having had clinical symptoms. In prospective studies in which participants are questioned about symptoms of PID, few cases are likely to be missed. The risk of *over*diagnosis of PID depends on the definitions used. The definition used in the POPI trial[19] corresponded very closely with 'probable/definite' PID as defined in *Chapter 2*. To the extent that the POPI trial[19] has a dominating influence on the pooled estimate, it is reasonable to assume that the estimate of progression risk applies to 'probable/definite' PID.

However, as noted in *Chapter 2*, although it is believed that a high proportion, 66% of PID is undiagnosed, only a small proportion, 11%, is asymptomatic.[103] In the POPI trial[19] women were alerted to the risks of PID as part of the consenting process, as follows: '. . . This is why if you could have been at risk of infection, or have any symptoms which could be due to a sexually transmitted infection, it is vital that you have a check up at a genitourinary clinic even if you are participating in the study'.[221] Women would also have been aware that they would be followed up at the end of the study. Our assumption therefore is that in the POPI trial,[19] and probably in all prospective studies following women at risk of STIs and PID, only the small proportion of PID – the 'silent' proportion – will be missed.

The progression risk from CT to PID was estimated to be 14.8%. This estimate is based on the trial data alone. A previously published rate from this study,[201] which was a little higher (16%), was based on the trial together with the single non-randomised study with a control group. However, in discussing the progression rate from CT to PID, we should bear in mind not one single estimate but three estimates, each reflecting a different level of ascertainment of PID. The 14.8% estimate most likely represents the 87% of PID that is symptomatic, a 17.1% estimate would be more appropriate for *all* PID, diagnosed and undiagnosed, whereas an estimate of approximately 6% would correspond to the level of ascertainment that is represented in routine statistics.

The systematic difference in observed PID rates in studies with shorter and longer follow-up periods has led to the suggestion that the progression rate is not constant over time.[98] The potential impact of this has been explored in the mathematical modelling literature, using Markov or similar models[208,222] to predict the proportion of PID prevented. However, with one exception,[222] these studies have generated estimates of the proportion prevented from scenario analyses based on estimates of progression rates that were derived from models that do not allow for the competing risk of clearance or the left truncation inherent in screening studies. The key contribution of this paper is to actually estimate all of these parameters consistently within the Markov model framework. We found that the proportion of CT-related PID that

could be prevented by annual screening may be as low as 55%, which might represent only 20–25% of *all* PID in younger women[62] (see later chapters). Estimates based on two-rate models were relatively insensitive to the duration of a period of heightened PID risk following CT infection, but were derived from a single, non-randomised study. The sparse data available do not distinguish homogeneous and piecewise homogeneous models statistically, and the credible intervals on the parameters are wide. However, both one- and two-rate models are considered plausible a priori.[98]

Summary of assumptions and findings

Assumptions

1. Studies based on screening asymptomatic individuals recruit *prevalent* infection, whereas clinical-based studies recruit recently *incident* cases.
2. The definitions of PID used in the POPI study[19] correspond closely with the definition of probable/definite PID used in *Chapter 2*.

Summary of findings

1. The data do not distinguish between a model that assumes a constant hazard of PID following CT infection, and a model in which the rate is higher over the initial 1–3 months.
2. The results of the POPI trial[19] are consistent with the results of other trials and controlled prospective studies although there is a high degree of uncertainty.
3. The CT-related risk of PID following CT infection is 14.8% (95% CrI 4.8% to 24.8%), based on the trial evidence on PID that would be ascertained in a prospective study.
4. The total CT-related risk of PID following CT, including undiagnosed PID, is 17.1% (95% CrI 5.6% to 28.9%).
5. 61% (95% CrI 55% to 67%) of CT-related PID could be prevented by screening at annual intervals. The estimate is slightly lower if CT-to-PID progression rates are higher in the first 1–3 months.

Chapter 7 Chlamydia and pelvic inflammatory disease: Population Excess Fraction based on prospective, retrospective and routine data

Objectives

To:

1. generate and compare estimates of the incidence of PID
2. generate and compare plausible estimates of the proportion of PID attributable to CT (the PEF)
3. produce coherent estimates of: CT incidence, PID incidence, the CT-to-PID progression rate, and the PEF.

Introduction

The previous chapter produced estimates of the risk of clinical PID (as defined in *Chapter 2*) caused by an episode of CT, based on prospective, and, particularly, randomised, data. We now consider the relation between CT and PID from a broader perspective: not only the risk of PID following CT, but also how much PID is attributable to CT. Viewed in isolation, the two questions are technically independent of each other. However, suppose we had independent estimates of (1) CT incidence; (2) PID incidence; (3) the CT-to-PID progression risk; and (4) the proportion of PID attributable to CT, the PEF. It would then be possible to generate a prediction for each of these quantities from the other three. Therefore, estimates of the two quantities are *not* independent if we examine the available data sources more broadly.

In fact, the relation between untreated CT and PID has been studied in a wide variety of ways, drawing on prospective, retrospective and routine data. Following a MPES approach, we want to know whether the different methods give consistent answers.

Given that previous chapters have dealt with CT incidence and the CT-to-PID progression risk, in order to provide a complete account of CT and PID, we must answer two remaining questions: first, what is the incidence of all-cause PID; and second, how much PID is caused by CT. We begin by describing evidence sources and statistical methodology for each topic. Two sources of evidence of PID incidence are considered: a direct estimate from the control arm of the POPI trial[19] and an estimate derived from routine data. We then consider five separate methods of estimating the PEF in turn. For each estimate we first review the evidence sources available, describing the relationships between the different elements. We then set out the statistical models used to obtain estimates from the data. In the results we assess the consistency of evidence on PID incidence, and compare the various estimates of PEF. In discussion we review some of the key assumptions made and provide a rationale for our choice of PEF estimate to take forward to later chapters. We also present a coherent set of estimates of CT incidence, prevalence and duration, CT-to-PID progression rate, PID incidence and the PEF.

We note here that there is a wider literature on the relation between CT and PID which we have not reviewed here, although it is, on the face of it, highly relevant. In particular, we have not included estimates of the CT-to-PID progression risk based on register studies[36,215] because women with CT in these studies are treated, and are therefore not able to inform the CT-to-PID progression risk in untreated infection. Further, although these studies have been used for this purpose in the past,[2,15] the CDC Task Force[98] also concluded that such studies could not help determine the relationship between CT and PID.

A second set of estimates not reviewed here are those of van Valkengoed *et al.*,[39] who used a complex procedure to estimate the prospective risk of PID, EP and TFI, following CT, from retrospective and routine data sources. These risk estimates are extraordinarily low, and a methodological critique is given in *Appendix 9*.

Methods: review of evidence sources and statistical estimation

Estimates of the incidence of all-cause pelvic inflammatory disease

This section describes two independent sources of PID incidence estimates, and a third estimate based on pooling them.

Estimates based on the control arm of the Prevention of Pelvic Infection trial

A comparatively direct estimate of $\lambda_a^{ALL\ PID}$ can be derived from the control arm of the POPI trial,[19] if we assume that the trial sample is approximately representative of the general female population of the same age, approximately age 16–24 years. In the unscreened arm, 23 (r^{POPI}) cases of clinical PID were reported in a sample of 1186 (n^{POPI}) women, aged 16–27 years, followed up for a period of 1 year. Following comments on the definition and ascertainment of PID in *Chapters 2* and *6*, we assume that all *symptomatic* PID meeting the 'probable/definite' criteria will be ascertained, including those normally undiagnosed. The POPI estimates are therefore 'grossed up' to account for the underascertainment of asymptomatic (silent) PID in prospective studies. The POPI[19] control arm estimates $\lambda_a^{ALL\ PID}.\psi^{Sym}$, where ψ^{Sym} is the proportion of PID that is symptomatic.

$$\begin{aligned} r^{POPI} &\sim Bin(p^{POPI}, n^{POPI}) \\ p^{POPI} &= 1-\exp(-\lambda^{POPI}) \\ \lambda^{POPI} &= \lambda^{All\ PID}.\psi^{Sym} \\ r^{Sym} &\sim Bin(\psi^{Sym}, n^{Sym}) \end{aligned} \tag{22}$$

The probability that incident PID is symptomatic ψ^{Sym} is informed by the 32 out of 36 cases who reported symptoms in the Wolner-Hanssen study.[103]

Incidence of all-cause pelvic inflammatory disease based on routine data

A second, more indirect, estimate of incidence can be derived from PID data routinely collected in England. These data represent only diagnosed 'probable/definite' PID, and so must be grossed up by an estimate of the proportion of PID that is diagnosed in a general population context (see below).

There are three sources of data on PID in England: HES,[117] GPRD[90] and KC-60 (the routine returns from STI clinics[198]) (*Table 18*). The three sources pick up cases from different care pathways. HES[117] reports the number of cases of PID diagnosed in hospital (codes N70-N74). GPRD[90] reports all diagnosis made in an

TABLE 18 Number of incident cases of PID in England, 2002

Age (years)	HES (A)	GPRD[a] (B)	GUM clinics[b] (C)	MinC + Max (B,A)	Max A + B + C	Female population
16–19	1233	5083	3212	8295	9528	1,199,600
20–24	3101	8842	4399	13,241	16,342	1,519,100
25–34	9756	14,932	3919	18,851	28,607	3,502,100
35–44	10,526	9609	1388	11,914	21,523	3,795,600

Max, maximum; Min, minimum.
a Definite and probable PID as defined in French *et al.*[90]
b Data by age not available for 2002 so we assume that the age distribution for these data were the same as in 2009.

approximately 8% sample of patients registered with general practices across the country. These data do not include PID as a separate diagnosis, so we take our estimate from a recent paper that used the recorded clinical descriptions of symptoms to estimate population incidence of definite and probable, and of possible, PID diagnosed in this setting.[90] Finally, PID diagnosed in STI clinics is reported as 'complicated chlamydia'. Data shown in *Table 18* is for 2002, the final year before the introduction of the National Chlamydia Screening Programme.

To use these data to derive age group-specific estimates of the annual incidence of diagnosed PID, we need to make assumptions about the degree of overlap between the three sources. Consultation with experts suggests that referrals of PID cases to and from departments of GUM (STI clinics) were rare prior to 2008. Referrals from GP practices to hospital are likely to have been more common for severe cases. A primary care guideline for STI management published in 2006[223] recommended referral to a GUM clinic for mild/moderate cases of PID, and immediate commencement of treatment was advised if an urgent appointment was not available. Prior to 2008, access to GUM services within 48 hours was poor.[105] It was recommended that severe cases should be referred to gynaecology for admission to hospital. We therefore assume that the total of the STI, GPRD and HES data, within each age group, represents an upper bound for the number of PID cases diagnosed in England each year. A minimum was formed by adding the number of GUM cases to either the GPRD or the HES cases, whichever was largest.

Routine data on diagnosed PID, given information on age-specific population sizes N_a, provides information on $\lambda_a^{DiagPID}$. An estimate of total (all cause) rate of PID can then be constructed by using information on the proportion diagnosed, ψ^{Diag}:

$$\lambda_a^{ALLPID} = \frac{\lambda_a^{DiagPID}}{\psi^{Diag}} \tag{23}$$

Binomially distributed data based on *Table 18* are used to inform minimum and maximum limits for the annual probability of (diagnosed) PID, and these are used to define the upper and lower limits for $\lambda_a^{DiagPID}$. The WinBUGS 'cut function' was used to ensure that this uniform distribution was not updated (see *Chapter 3*).

$$\begin{aligned}
&r_a^{min} \sim Binomial(p_a^{min}, N_a), \quad r_a^{max} \sim Binomial(p_a^{max}, N_a)\\
&p_a \sim Uniform(p_a^{min}, p_a^{max})\\
&\lambda_a^{DiagPID} = -\log(1-p_a)\\
&\lambda_a^{ALLPID} = \lambda_a^{DiagPID}/\psi^{Diag}\\
&r^{Diag} \sim Bin(\psi^{Diag}, n^{Diag})
\end{aligned} \tag{24}$$

where N_a is the number of women in each age group in England in 2002 from Census data.

The degree of underdiagnosis was based on a single study of 36 infertile women, only 11 of whom (30.6%) reported that they had been diagnosed with PID.[103] This study reported data with a binomial likelihood.

Synthesis of pelvic inflammatory disease incidence estimates

The two estimates of PID incidence were compared. A DAG (*Figure 14*) shows how the two estimates were combined. The WinBUGS code and data sets are set out in *Appendix 10*.

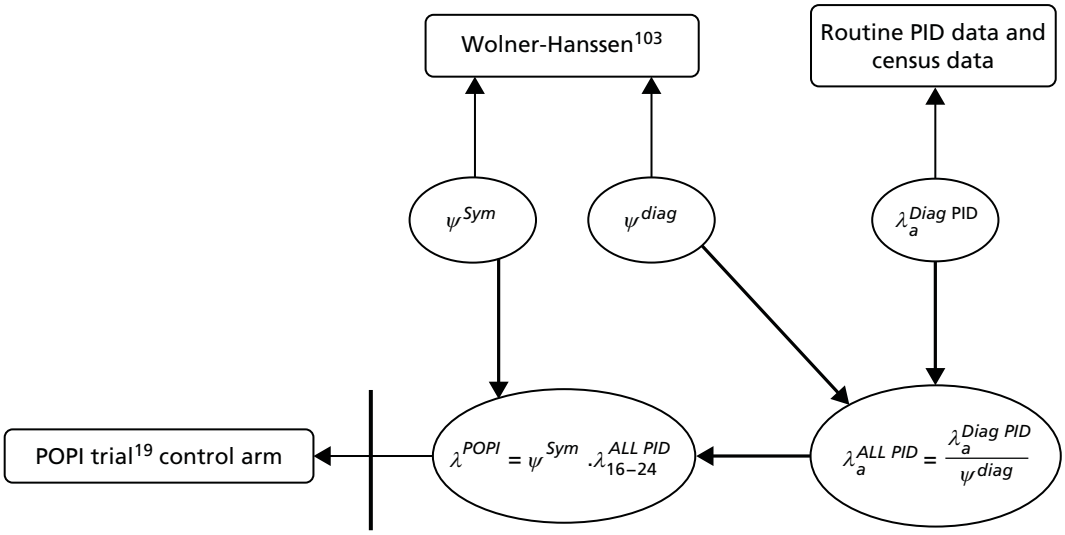

FIGURE 14 Directed acyclic graph showing combination of two sources of evidence on PID incidence.

Population Excess Fraction: how much PID is caused by chlamydia?

In this section we examine estimates of the PEF based on microbiological studies, estimates based on retrospective case–control studies, and estimates derived from the ratio of CT-related PID incidence and all-cause PID incidence, and finally estimates based on randomised trials. DAGs for each of the estimates are shown in *Figure 15*.

Uncontrolled studies of the microbial aetiology of PID

Microbiological analysis of samples from the genital tract of women with PID provides direct information on potentially causative organisms. Based on 19 microbiological studies conducted between 1977 and 1992, in a number of countries, Paavonen *et al.*[82] reported that CT was involved in 30% of PID cases. It has been argued[20] that this must be an underestimate of the current role of CT, as gonorrhoea was a common cause of PID during the period when many of these studies were undertaken, and it is generally agreed that gonorrhoea is a far less common aetiology at present, particularly in Europe.

Simms and Stephenson[62] provide a summary of studies in which patients with laparoscopically proven PID, in other words salpingitis (*see Chapter 2*), recruited in various clinical settings, were examined for evidence of current CT infection. The proportion with evidence of current CT in upper genital tract samples varied from 12% to 65%, reflecting considerable variation over time and between countries. The largest UK study,[224] conducted between 1989 and 1993, reported 39%. In another UK study,[37] conducted from 2000 to 2002, 42 of 140 (30%) salpingitis cases had evidence of exposure to CT.[37]

It is tempting to consider these studies as producing a direct estimate of the proportion of PID that is attributable to CT.[1] But this may be misleading for several reasons, the most obvious of which is that a control group is required consisting of women without PID. The proportion of control samples in which we would expect to find evidence of CT is, of course, quite small. The general population prevalence of CT is in the order of only 3% (see *Chapter 5*).

A more serious weakness is that the results obtained depend markedly on the sites from which samples are taken. In particular, the presence of CT in the upper genital tract, which is more likely to be causally related to PID, may not be well predicted by its presence in the lower genital tract.[225,226] These issues were clarified by a recent study in Erfurt, Germany,[225] looking at 363 women with laparoscopically confirmed PID. CT was found in the genital tract of 103 (28.4%), and in 55 (15.2%) it could be isolated from the cervix. In 23 (6.3%), CT was isolated from both the cervix and the fallopian tubes, whereas in 47 (12.9%) CT was isolated in the fallopian tubes *only*.

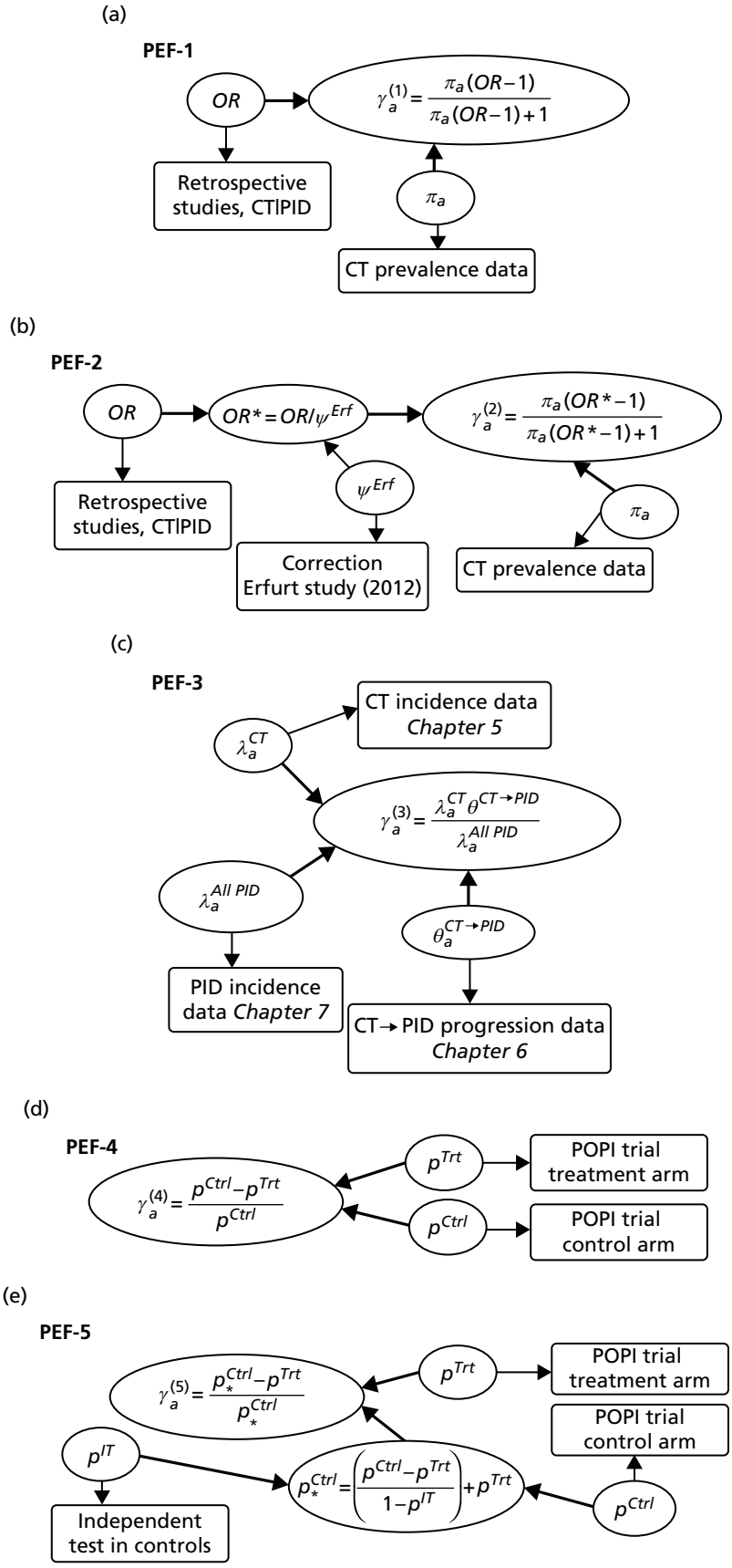

FIGURE 15 Directed acyclic graph showing methods of estimating Population Excess Fraction. Symbols: PEF-1: π_a, CT prevalence in age band *a*; OR from case–control study. PEF-2 OR* adjusted OR; ψ^{Erf} correct for under-ascertainment of CT infection based on Erfurt study.[225] PEF-3: λ_a^{CT} incidence of CT; $\lambda_a^{ALL\ PID}$ incidence of all-cause PID; $\theta_a^{CT\rightarrow PID}$ CT-to-PID progression risk. PEF-4: P^{trt} probability of PID in POPI[19] treatment arm; p^{Ctrl} probability of PID in POPI[19] deferred treatment (control) arm; p_*^{Ctrl} adjusted p^{Ctrl}; p^{IT} probability independent testing in CT+ controls.

If we take infection in the fallopian tubes as the causal factor in subsequent reproductive damage, this implies that studies based on cervical isolates may be underestimating the role of CT in PID-related reproductive damage by a factor of 103/(103–47) = 1.84. The Erfurt study[225] found CT in the fallopian tubes of just under 20% of cases of laparoscopically confirmed PID. Few studies of PID routinely sample *both* the cervix and fallopian tubes for CT, so direct comparisons are difficult. However, high titre antibody can be used as a surrogate for upper genital tract infection.[85] Taylor-Robinson *et al.*[88] also observed that infection at the cervix probably underestimates the role of CT in PID-related reproductive damage by a similar factor 16/(16–6), or 1.6. This was based on the observation that of the 22 women diagnosed with acute salpingitis on laparoscopy: 10 had CT detected at the cervix and an *additional* six had high-titre serum CT IgG antibody.

These considerations help to quantify the likely degree of underestimation of the OR, and hence the PEF, from case–control studies based on lower genital tract samples, and this can be used to adjust estimates from those studies.

Population Excess Fraction based on case–control studies

We noted in *Chapter 3* that the presence of positive confounding in estimates of the RR of the disease (PID) given the exposure (CT), if based on observational data, would tend to result in *overestimation* of the PEF. This can be mitigated to some extent by choosing study designs in which the risk of confounding is minimised. For example, it is preferable to use a marker of *current* CT infection rather than a marker of previous CT infection. We have therefore used data only from case–control studies that use current infection as a marker of exposure. (We set aside, here, the further difficulties with the serological markers used in the published literature, namely their poor sensitivity and specificity: this is discussed at greater length in *Chapter 11*.)

Identification of case–control studies on CT prevalence in women with PID and control subjects

We searched for studies using current CT infection as a marker of CT exposure. Strictly speaking, in a synthesis focusing on prevalence and sequelae of CT in the UK, only contemporary UK data should be used. We have, nevertheless, used data from Europe on the basis that the epidemiology of STIs is generally similar in Western European countries. Studies from North America have been excluded because gonorrhoea has tended to have a more important role in the aetiology of PID in North America than the UK.[227–229] We also excluded studies published before the 1990s. Our literature identification process (see *Chapter 3*) identified three studies,[31,37,230] shown in *Table 19*.

TABLE 19 Odds ratios from retrospective studies

Study	Group	Data[a]	Crude CT prevalence	Crude OR
Paavonen[31]	Cases	13/30	0.43 (0.28 to 0.63)	6.9 (0.77 to 61.4)
	Controls[b]	1/10	0.10 (0.03 to 0.44)	
Mascellino[230]	Cases	22/110	0.20 (0.14 to 0.29)	7.0 (3.1 to 15.8)
	Controls	9/261	0.03 (0.02 to 0.06)	
Simms[37]	Cases	17/140	0.12 (0.08 to 0.19)	18.7 (2.45 to 142)
	Controls[b,c]	1/136	0.01 (0.00 to 0.04)	

a Binomial numerators and denominators.
b Illustrative, as there are insufficient numbers to assume asymptotic normality.
c GP control Group used.

Estimates of the Population Excess Fraction from case–control studies, PEF-1 and PEF-2

A pooled OR was estimated from these studies (see below), and used to derive an estimate of the PEF, using equation 1. Although the same OR is assumed to apply to women of all ages, the formula will generate age-specific estimates of PEF, because age-specific estimates of CT prevalence from *Chapter 5* are entered, as follows:

$$PEF_a^{(1)} = \gamma_a^{(1)} = \frac{\pi_a.(RR-1)}{\pi_a.(RR-1)+1} \tag{25}$$

The six data points are used to estimate four parameters: three study-specific 'baselines', the log odds in the control groups μ_s, with s indexing study, and one 'FE' log OR β. Using a standard logistic regression model, with 0 for controls, 1 for PID cases:

$$\begin{aligned} \text{logit}(\lambda_{s,0}) &= \mu_s \\ \text{logit}(\lambda_{s,1}) &= \mu_s + \beta \end{aligned} \tag{26}$$

The OR can be recovered via: $OR = \exp(\beta)$, and used as a RR in (see equation 25) (see *Chapter 3*) to generate a set of estimates, PEF-1.

A second set of estimates (PEF-2) were generated by adjusting the OR from the retrospective studies for under-ascertainment of CT in the PID cases, using the 1.84 estimate from the Erfurt study.[225] The uncertainty in the adjustment is taken into account by setting: $OR^{Adj} = OR/\psi^{Erf}$, where $\psi^{Erf} \sim Beta(56,47)$.

Estimates of Population Excess Fraction based on CT-related and all-cause PID incidence, PEF-3

Adams *et al.*[1] obtained estimates of all-cause PID incidence in 16- to 44-year-olds from a GP-based study.[231] Based on a retrospective study,[37] they reckoned that a maximum of 30% of PID could be attributed to CT, and used this to identify a range of estimates of CT-related PID incidence. They then multiplied a number of CT-to-PID progression rates into estimates of CT incidence, to identify progression rates that were consistent with their results on CT-related PID. We mention this study,[231] not to review the findings in detail but simply to draw attention to the general method, which we use below in an inverted form to estimate the PEF, although with different data inputs.

Here, the age-specific PEF is taken to be the ratio of the incidence of CT-related PID to the incidence of all-cause PID, where the incidence of CT-related PID is the product of CT incidence (see *Chapter 5*) and the CT-to-PID progression rate (see *Chapter 6*):

$$PEF_a^{(3)} = \gamma_a^{(3)} = \frac{\lambda_a^{CT}\theta^{CT\to PID}}{\lambda_a^{ALLPID}} \tag{27}$$

To implement this, we pool two independent sources for estimating all-cause PID incidence (see below), and estimates of CT incidence and CT-to-PID progression rate from *Chapter 6*.

For these last two parameters, we use a multivariate normal approximation to the posterior distribution estimated in the incidence, prevalence, duration synthesis (see *Chapter 5*), and this is entered as data in the calculations. Specifically, incidence and duration were normal on the log scale, the proportion symptomatic was normal on the natural scale, and prevalence was normal on a logit scale. Similarly, a normal likelihood for CT-to-PID progression risk is derived from the posterior analysis in *Chapter 6*.

Estimates of Population Excess Fraction based on screening trials, PEF-4 and PEF-5

The Scholes trial[13] compared PID risk 1 year later in women randomised to screening and treatment and women randomised to no screening. The observed RR was 0.44 (95% CI 0.20 to 0.88). This represents a RRR of 0.56, implying that 56% of PID is due to CT in the study population, and this can be interpreted as an estimate of the PEF (see *Chapter 3*). One of the early criticisms of the trial[18] was that this was an unrealistically optimistic result because it conflicted with the microbiological studies cited above, apparently showing that 30% of PID was CT related. Gottlieb *et al.*,[226] in an important review, also draw attention to this apparent discrepancy, but points out that a similar RR was observed in the Ostergaard trial.[12]

An approximate PEF can be derived from the RR of PID in the POPI trial, using equation 3.[12] There were 23 PIDs in the untreated group of 1186, and 15 in the treated group of 1191. An estimate of the PEF is therefore:

$$p^{Ctrl} \sim Beta(23,1163), \quad p^{Trt} \sim Beta(15,1176)$$
$$PEF^{(4)} = \gamma_a^{(4)} = \frac{p^{Ctrl} - p^{Trt}}{p^{Ctrl}} \tag{28}$$

However, this estimate is likely to be biased because 29 out of 67 (43%) untreated controls, who were initially CT+ undertook to be tested independently. We can assume, therefore, that the excess risk of PID in the untreated controls, $p^{Ctrl} - p^{Trt}$ has been underestimated by a factor of $1 - p^{IT}$, where p^{IT} is the probability that a CT+ control is independently tested and treated. Accordingly, we can form an adjusted estimate:

$$p^{IT} \sim Beta(29,38)$$
$$p_*^{Ctrl} = \left(\frac{p^{Ctrl} - p^{Trt}}{1 - p^{IT}}\right) + p^{TrT}$$
$$PEF^{(5)} = \gamma_a^{(5)} = \frac{p_*^{Ctrl} - p^{Trt}}{p_*^{Ctrl}} \tag{29}$$

This should be regarded as a 'maximum' correction, as it assumes that all the independent testing took place at the outset.

Results

Consistency of evidence on pelvic inflammatory disease incidence

Table 20 provides a series of age-specific estimates for the incidence of (all cause) PID, from the different data sources. The 'direct' estimate from the POPI trial,[19] 2.0% per year – when adjusted for under-ascertainment due to asymptomatic presentation – represents a 2.4% per year incidence. This is consistent with the estimate derived from routine data, 3.0% per year for that age range, following adjustment for the underdiagnosis inherent in routine PID statistics. *Figure 16* shows the posterior distributions of PID incidence from the two sources, and the pooled estimate, and establishes the consistency of the estimates. The consistency of registry data with the POPI study[19] can be studied only for the 16- to 24-year-old age group.

Note that when combining the direct and indirect evidence sources, the incidence information back-propagates to 'update' the estimate of the proportions of PID that is undiagnosed. This was 31% in the original study but is estimated to be 36% following synthesis with the POPI[19] data.

The final pooled PID incidence estimates are 2.5% per year (95% CrI 1.8% to 3.4%) in women aged 16–24 years, and 1.6% (95% CrI 1.1% to 2.2%) in women aged 25–44 years. Taking into account the age structure of the female population, we can estimate that 62.9% (95% CrI 57.8% to 67.4%) of PID episodes occur in women over 24 years.

TABLE 20 Results of synthesis of evidence on the incidence of all-cause PID in England

Parameters	Age (years)	POPI, unadjusted	Routine data, unadjusted	POPI, adjusted for under-ascertainment	Routine adjusted for underdiagnosis	Wolner-Hanssen	Synthesis model
PID incidence, % per year	16–19	–	0.74 (0.69 to 0.80)	–	2.6 (1.6 to 4.3)	–	2.1 (1.5 to 2.9)
	20–24	–	0.98 (0.88 to 1.1)	–	3.4 (2.1 to 5.6)	–	2.8 (2.0 to 2.8)
	25–34	–	0.68 (0.55 to 0.81)	–	2.3 (1.4 to 4.0)	–	1.9 (1.3 to 2.8)
	35–44	–	0.44 (0.32 to 0.56)	–	1.5 (0.83 to 2.7)	–	1.3 (0.78 to 1.9)
	16–24	2.0 (1.3 to 2.9)	0.88 (0.81 to 0.94)	2.4 (1.5 to 3.4)	3.0 (1.9 to 5.0)	–	2.5 (1.8 to 3.4)
	25–44	–	0.56 (0.45 to 0.66)	–	1.9 (1.1 to 3.3)	–	1.6 (1.1 to 2.2)
	16–44	–	0.64 (0.56 to 0.72)	–	2.2 (1.4 to 3.7)	–	1.8 (1.3 to 2.5)
% of PID diagnosed	16–44	–	–	–	–	30.8 (17 to 46)	36.0 (26 to 48)
% of symptomatic PID undiagnosed	16–44	–	–	–	–	56.4 (41 to 71)	51.4 (39 to 63)
% of PID asymptomatic	16–44	–	–	–	–	12.8 (4.4 to 25)	12.6 (4.3 to 25)
Residual deviance (data points)	–	1 (1)	4 (4)	1 (1)	5.9 (6)	1.9 (2)	6.7 (6)

Although consistency of registry data with the POPI study[19] can be checked in only the 16–24 years age group, estimates of PID incidence $\lambda_a^{ALL\ PID}$ for other age ranges can be derived using equation 23.
Posterior means and 95% CrIs.

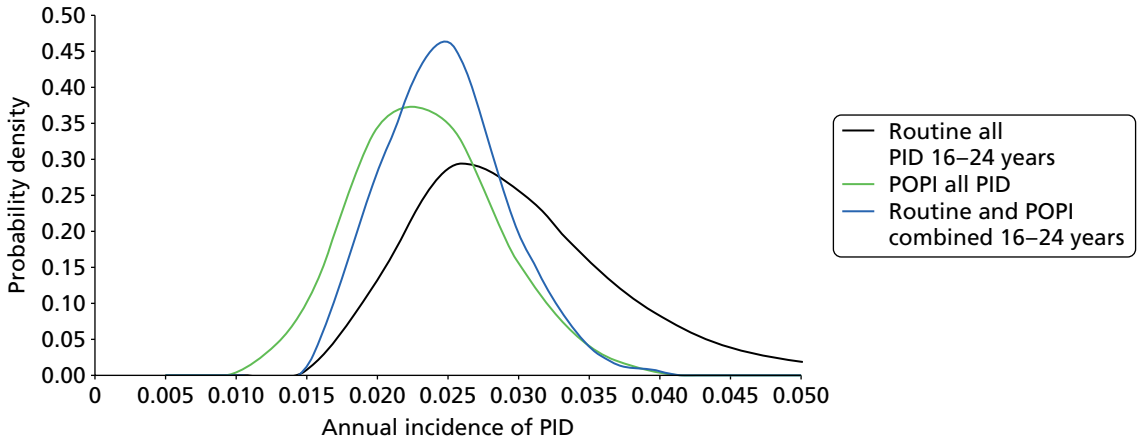

FIGURE 16 Evidence consistency plot for the incidence of all-cause PID in England in women aged 16–24 years.

Estimates of the Population Excess Fraction

The synthesis of the three retrospective studies[31,37,230] of CT in women with PID and control subjects (see *Table 19*) produces a posterior mean OR of 9.2 (95% CrI 4.4 to 18.1). The model fitted well (residual deviance of 5, compared with six data points). The OR adjusted by the under-ascertainment factor from the Erfurt study[225] is 17.1 (95% CrI 7.9 to 34).

The several estimates of the PEF are shown in *Table 21*. The adjustment for under-detection of CT infection in the case–control studies, based on the Erfurt study,[225] almost doubles the overall PEF (women aged 16–44 years) from 14.5% to 24.6%. The dramatic fall in PEF with age is no more than a reflection of our assumption that the OR is not related to age, whereas the incidence and prevalence of PID varies, as is evident from the routine data (see *Table 18*).

TABLE 21 Alternative estimates of the PEF, the percentage of PID caused by CT

Age (years)	Retrospective (OR = 9.2): case–control studies, see *Chapter 6*	Adjusted retrospective OR = 17.1, based on Erfurt study[225]	CT-related PID divided by all-cause PID incidence	POPI trial[19] alone, from reported RR	POPI trial,[19] adjusted for treatment during follow-up in the control group
	1	2	3	4	
16–19	33.8 (16.9 to 54.8)	49.5 (29.4 to 70.3)	53.5 (15.6 to 100)	–	–
20–24	23.7 (11.1 to 41.4)	37.4 (20.4 to 58.0)	24.3 (7.2 to 47.6)	–	–
25–34	8.5 (3.4 to 17.2)	15.2 (6.7 to 28.8)	10.3 (2.9 to 21.2)	–	–
35–44	6.2 (2.3 to 13.5)	11.5 (4.6 to 23.3)	11.5 (3.0 to 25.7)	–	–
16–24	28.6 (14.0 to 47.9)	43.5 (25.0 to 23.3)	35.3 (10.5 to 68.5)	35.7 (0 to 67.2)	49.4 (0 to 78.9)
25–44	7.3 (2.9 to 15.1)	13.3 (5.8 to 25.8)	10.6 (3.0 to 21.9)	–	–
16–44	14.5 (6.4 to 27.1)	24.6 (12.3 to 42.0)	19.7 (5.9 to 38.1)	–	–

Column 1: Estimates based on three case–control studies; column 2, same, but OR adjusted by data from the Erfurt study[225] (see text); column 3, ratio of CT-related PID incidence, estimated as the product of CT incidence (see *Chapter 5*) and CT-to-PID progression risk (see *Chapter 6*) to all-cause PID incidence (synthesis model, see *Table 20*); column 4, median PEF derived directly from RR reported in the POPI trial;[19] and column 5, median PEF from POPI trial[19] adjusted for independent testing in the deferred treatment group.

The PEF-3 estimates (column 3) based on the ratio of CT-related PID to all-cause PID also show a marked falling off with age. It is reassuring that the adjusted estimates based on case–control studies are close to estimates derived in a very different way from the progression rate, CT incidence and all-cause PID incidence.

The direct estimate of the PEF based on the POPI trial,[19] PEF-4, is 35.7%, but if this is corrected for independent testing in the deferred treatment (control) arm, this becomes 49.4%. The CrIs on both estimates are wide, and span zero because the difference in the crude PID rates between the two trial arms itself is too uncertain to rule out no difference with 95% confidence.

Discussion

The chapter establishes that routine data on PID incidence collected from GP, hospital and STI clinic sources, taking account of the likely overlap between these sources and the extent of undiagnosed PID, is consistent with an estimate of PID derived from the POPI[19] trial.

Second, setting aside the actual estimates obtained, we have established for the first time a coherent framework in which CT incidence and prevalence, PID incidence, CT-to-PID progression rate, and the PEF can all be considered together. The coherent set of estimates obtained, for the first time in any country or region, reflect what we feel is the totality of evidence on all these parameters, subject, of course, to our interpretation of the evidence sources.

Although one might wish for greater precision in each one of the data inputs that have been examined here, we would suggest that the analysis lends a degree of robustness to our interpretation of all of the data sources, but particularly to our quantitative information on the causal connection between incident CT and PID. This estimate of the causal connection between CT and PID is pivotal, as the further downstream sequelae of CT, namely EP and infertility, are viewed as arising from CT-related PID, and not from CT in the absence of PID.

The general consistency between the various estimates of PEF, based on fundamentally different sources, is reassuring, although the CrIs are wide. The two direct estimates based on the POPI trial,[19] PEF-4 and PEF-5, probably represent lower and upper limits, and they are close to estimates from previous trials. Interestingly, of the 26 PID patients in the POPI trial[19] who were tested, 16 (61.5%) were CT+ at the time the PID was diagnosed. Of the 16 who were CT– at the start of the trial, 10 (62.5%) were CT+ at diagnosis. These further observations from the POPI trial[19] suggest that that the corrected estimate may be closer to the truth.

Although the failure to find inconsistency raises the credibility of the analyses, confidence in the estimates has to be tempered by their relatively low precision, which makes inconsistency difficult to demonstrate. Moreover, each of the analyses are subject to a number of caveats.

Limitations of the analysis

Overlap between sources of routine data on incident pelvic inflammatory disease

One important limitation is the poor information currently available on the degree of overlap between the three routine data sources contributing information on incident PID (see *Table 18*). Our assumption was that the true number of incident cases was from a uniform distribution bounded by a ‘minimum’ (GUM plus the largest of GPRD and HES) and a ‘maximum’ in which there was no overlap. Although this seems a reasonable way to use the data available, the ‘gap’ between the minimum and maximum estimates was 30–37% of the lower estimate, depending on age. A recent paper on PID management based on GPRD noted that only 56% of the 3669 cases examined were managed entirely within the practice. If the age-specific GPRD figures are divided by 0.56, the number obtained lies between the minimum and maximum, which provides some support for the procedure we have followed.[97]

Obtaining a more precise estimate of the number of incident cases, by collecting data on referrals between GPRD, hospital and GUM clinics, could improve the accuracy of estimates considerably. This would be a useful avenue for further research (see *Chapter 12*).

Proportion of pelvic inflammatory disease patients diagnosed and ascertained

Our estimate of the proportion of PID episodes diagnosed is based on a single study looking at infertile women with evidence of past PID.[103] This study[103] has informed two key parameters which have been used to 'gross up' each of two estimates of PID incidence. First, we grossed up the estimate seen in the POPI trial,[19] to account for the proportion of 'silent' PID that would not be observed even in a prospective study. Second, we use it to gross up evidence from routine data, to account for both silent PID and for symptomatic PID that does not lead to medical treatment or which is not diagnosed. This single, small study,[103] carried out in another country 30 years ago, is being asked to carry quite a considerable weight and this must be recognised as a limitation of our analyses.

Retrospective data and confounding

As noted earlier in this chapter and elsewhere, the use of equation 25 to estimate the PEF from retrospective data is inherently likely to overestimate it as a result of positive confounding with other exposures, both other STIs that are known to be causal agents for PID, and non-STI infections that can be transmitted as a result of sexual intercourse. It is therefore possible that the PEF-1 estimate, and particularly the PEF-2 estimate, which is adjusted for probable under-ascertainment of CT infection in women with PID, represent overestimates. On the other hand, this adjustment may not fully account for under-ascertainment if a proportion of women with CT-related PID may clear their infection before the diagnosis of PID is made.[88,226] In addition, estimates of PEF are time- and place-dependent: the extent to which we can generalise from these results to the UK situation is very hard to assess. Therefore, although the PEF-1 and PEF-2 estimates based on retrospective data are of interest, we conclude that they are biased to an extent that is had to quantify, and they play no further role in this report.

Impact of age on the relationship between Chlamydia trachomatis *and pelvic inflammatory disease*

The PEF estimates show a clearly declining trend with age. In the retrospective estimates, this follows as a consequence of assuming a constant OR and applying this to a prevalence that declines steeply with age. In the case of PEF-3, the decline with age is partly due to the decline in CT incidence with age (exactly mirroring the decline in prevalence with age), the age profile of PID in the routine data, and the assumption that the proportion of PID that is undiagnosed is constant over age.

Although the decrease in PEF with age is to some extent a result of our assumption that neither the probability that PID is diagnosed nor the risk of PID following CT are age-dependent, the degree of variation with age, a factor of 4–6 between age 16–19 years and age 35–44 years, is so extreme that we would have to propose quite extreme trends in one or both of those quantities to reverse it. The proposal that PEF decreases with age is, apparently, not one that has been made before, but it is consistent with the observation that estimates based on the trial evidence tend to be higher because the trials are heavily weighted towards younger women.

Age dependency in PEF, if confirmed (see *Chapter 2* for discussion of alternative explanations), would have significant impact on the public health importance of CT, as the majority of EP and TFI occurs in older women. Although it is certainly possible that CT infections in younger women have a key role in reproductive health problems that emerge many years later, these results focus attention on the distinctly different age profiles of CT and PID. There are insufficient data in either the prospective or retrospective studies to address this question more fully. Further work to examine these issues directly is proposed in our research recommendations.

Post-hoc element in choice of estimates

We record here that, when the work in this chapter was first carried out, we included only the unadjusted (PEF-1), and we pooled that information with the evidence sources contributing to estimate PEF-3. At that time we believed that the estimates based on retrospective studies must be upper bounds, and we did not examine the trial-based estimates. After we had completed the work on TFI (see *Chapters 10* and *11*), we became aware that the PEF-1 estimates were too low to be consistent with the results of those chapters. We then reconsidered the relationship between CT and PID, and re-examined the trial-based estimates of PEF, particularly in the light of the comments in Gottlieb *et al.*[226] At the same time, results from the recent Erfurt study[225] made it clear that the retrospective estimates were not necessarily upper bounds. There is, therefore, a post-hoc element to the reasoning in this chapter. The 'hypothesis-generating' aspect to our findings on PEF are reflected in the research recommendations. Nevertheless, that fact that two estimates of PID incidence are consistent with each other, and with an account of the PEF that is consistent with other independent sources of data, to some extent mitigates the post-hoc process in our investigation.

Coherent estimates of CT and PID incidence, CT-to-PID risk, and Population Excess Fraction

There is an option to strengthen our estimate of the PEF by pooling PEF-3 with either the adjusted retrospective evidence PEF-2, or the evidence directly derived from the POPI trial,[19] PEF-4 or PEF-5, or both. We have chosen not to do this, on the basis that the PEF-2 estimate is vulnerable to many biases, and relates to a different time and place, as discussed above. The estimate of PEF from the POPI trial[19] is not statistically independent of the estimate of the CT-to-PID risk estimate, and is also subject to uncertainties described above. However, we regard the similarity of these independently derived estimates as an important validation.

In taking forward the PEF-3 estimate $\gamma_a^{CT \leftarrow PID}$, which is calculated from the CT incidence estimates in *Chapter 5*, the CT-to-PID progression risk in *Chapter 6*, and all-cause PID incidence in this chapter, we have generated a fully coherent set of estimates for CT incidence, prevalence, duration and the proportion of infection that is symptomatic; PID incidence, the proportion of PID that is symptomatic and the proportion that is diagnosed, the CT-to-PID progression risk, and the PEF. The summary posterior means and CrIs can be found in *Tables 13* (full synthesis model), *17* (trials only), *20* (synthesis model) and *21* (column 3).

These estimates are carried forward to the analyses in the subsequent chapters.

Summary of assumptions and findings

Summary of assumptions

1. The probability that PID is diagnosed is independent of age.
2. The risk of PID due to CT is independent of age.
3. The probability of diagnosis is the same in CT- and non-CT-related PID.
4. In prospective studies, including the screening trials, all PID – except for asymptomatic PID – is diagnosed.
5. In routinely collected data on PID, a lower proportion of all PID is diagnosed than in prospective studies.

Summary of findings

1. Two estimates of PID incidence in the UK were consistent: one derived from the control arm of the POPI trial,[19] and the other from routine statistics combined with evidence on the proportion of PID that is diagnosed.
2. A pooled estimate of PID incidence in age 16- to 24-year-olds is 2.5% per year (95% CrI 1.8 to 3.4), and 1.8% per year (95% CrI 1.3% to 2.5%) in 16- to 44-year-olds.
3. 36% (95% CrI 26% to 48%) of PID that causes reproductive damage is diagnosed; 12.6% (95% CrI 4 to 25) is asymptomatic.
4. Estimates of the PEF were compared: (1) an estimate based on retrospective studies adjusted for under-ascertainment; (2) an estimate based on CT incidence, CT-to-PID progression rate, and PID incidence; and (3) estimates based on the POPI[19] trial. There was no evidence of inconsistency although CrIs were wide.
5. We estimate that 62.9% of PID episodes (95% CrI 57.8% to 67.4%) occur in women aged over 24 years.
6. The PEF declines sharply with age.
7. Our estimate of PEF, derived from CT and PID incidence data and the CT-to-PID progression risk is 19.7% (95% CrI 5.9% to 38.1%) in women aged 16–44 years, and 35.3% (95% CrI 10.5% to 68.5%) in women aged 16–24 years.

The analyses thus far in the report establish a set of coherent estimates for the UK of CT incidence, CT prevalence, CT duration, probability that an incident CT infection is symptomatic, PID incidence, proportion of PID that is diagnosed, proportion of PID that is symptomatic, the CT-to-PID progression, and the PEF of PID due to CT.

Chapter 8 Cumulative incidence of pelvic inflammatory disease: results from a Markov model

Objectives

1. To estimate the proportion of women, by age, who have ever experienced (1) an episode of PID; (2) an episode of salpingitis; (3) an episode of non-CT-related PID; and (4) an episode of non-CT-related salpingitis.
2. To estimate the proportion of women, by age at index salpingitis who experience zero, one, or two or more subsequent episodes of PID.

Introduction

The critical evidence on the relation between PID and both EP and TFI derives from the Lund study, represented by a series of classic papers published between 1980 and 1992,[33,110,124,232] although new analyses of these data continue to appear. Named after a county in Sweden, the Lund study was based on the relatively long-term follow-up of large numbers of women hospitalised with a diagnosis of PID that was confirmed as salpingitis on laparoscopy. Control subjects with the same initial diagnosis but no evidence of salpingitis on laparoscopy were also followed up, with subsequent episodes of PID, EPs and TFI as the recorded outcomes. The findings are described in detail in *Chapters 9* and *10*, respectively.

The Lund study established three clear principles:

The risks of EP and TFI:

- were raised above the background rate *only* in women with salpingitis; there was no evidence of heightened risk following PID that was not confirmed as salpingitis on laparoscopy
- increased with the severity of the salpingitis
- increased with the number of subsequent episodes of PID.

In *Chapters 9* and *10* we derive predictions of the numbers of EP and TFI cases that would occur in the UK, based on the risks observed in the Lund study. To do this we need to estimate the proportion of women who have been exposed to salpingitis, and the proportion of these who have experienced zero, one, two or more subsequent episodes of PID. In this chapter we provide such estimates, based on the PID incidence estimates from *Chapter 7*, the evidence on repeat infections in CT from *Chapter 5*, and the data reported in the Lund study[33] on the proportion of women with PID who had experienced one, two, three or more episodes of PID subsequent to their index salpingitis.

It needs to be recognised that to achieve this analysis a number of assumptions have to be made, which cannot be properly supported by empirical evidence, which are even somewhat arbitrary, and which, at best, must be considered as no more than 'reasonable'. In the discussion section of this chapter we will nonetheless 'reality check' our results on the cumulative incidence of PID and salpingitis, and the prevalence of one, two or more previous PIDs, by comparing our predictions to other sources of evidence on PID incidence and cumulative incidence.

Methods

Model for repeat pelvic inflammatory disease episodes

Estimates of age-specific incidence of CT-related PID, all-cause PID, and therefore non-CT-related PID are available from earlier chapters. What we need, however, is a way to estimate what proportion of these PID episodes are occurring in women *who have already experienced a previous episode*. In a homogeneous population, PID incidence would be the same in women who had previously had PID and those who had not. However, in the Lund study[33] it was observed that the majority of women who experience a second PID did so within 2 years. We therefore assume that women acquire PID at a rate $\lambda_{a,1}^{PID}$ in age group a, and that this rate increases to $\lambda_{a,2}^{PID}$ for the 2 years following a PID. To determine the ratio of $\lambda_{a,2}^{PID}$ to $\lambda_{a,1}^{PID}$ we use information from the LaMontagne study[192] of CT infection and re-infection rates for up to a 18-month period. This was described in *Chapter 5*. Specifically, we use the ratio on the GP group, which was 7.08 (3.97 to 11.6) from *Table 13*, and assume that this same ratio applies to PID incidence so that: $\lambda_{a,2}^{PID} = \eta^{GP}.\lambda_{a,1}^{PID}$.

The GP group was chosen, rather than the FP or STI groups, as it was the most likely to reflect re-infection rates in the general population. Using data on recurrent CT makes the further strong assumption that the pattern of infection and re-infection in CT is the same as the pattern of PID and repeat PID, for PID from any cause. Apart from the observation by Westrom[33] that repeat PIDs usually occurred within 2 years of follow-up, this assumption, and its many implications, do not have strong a priori validity. Our results are, however, consistent with the Lund study[33] observations on the incidence of repeat PID, and with other data, as we show below.

We have posterior estimates of η^{GP} from *Chapter 4* and $\lambda_a^{ALL\ PID}$ from *Chapter 7*, and our task is to calculate $\lambda_{a,1}^{PID}$ and $\lambda_{a,2}^{PID}$. The parameter $\lambda_a^{ALL\ PID}$ denotes the incidence of all PID, whether first, second, or subsequent, and from any cause. There are several ways to parameterise the problem but the simplest is as follows:

$$\begin{aligned} \lambda_{a,1}^{PID} &= \rho.\lambda_a^{ALL\ PID} \\ \lambda_{a,2}^{PID} &= \lambda_{a,1}^{PID}.\eta^{GP} \end{aligned} \tag{30}$$

The equations, together with estimates of $\lambda_a^{ALL\ PID}$ and η^{GP}, uniquely determine ρ, which is a calibration constant ensuring that when we sum first PIDs and subsequent PIDs we end up with the correct total PID incidence. It is interpreted as the ratio of PID incidence in women who have not had a PID in the last 2 years to the incidence of PID in all women. ρ is calculated through a trial and error process. The model was run from ages 16–44 years for different values of ρ until a value was found where the estimated incidence of any PID agreed with our estimates of PID incidence from *Chapter 7* (see *Table 20*).

Proportion of clinical pelvic inflammatory disease confirmed on laparoscopy

The Lund study[33] analysis relates to hospital-diagnosed PID that has been laparoscopically confirmed. In the Lund study[33] itself, the 'control subjects' were women admitted with PID in whom there were no findings of inflammation on laparoscopy. Interestingly, the proportion of all those referred with clinical PID who were confirmed by laparoscopy fell systematically from 80% in those recruited 1960–4 and 1965–9, to 78% in 1970–4, to 70% in 1975–9, and to 60% in 1980–4.[110] A 2003 paper by Simms *et al.*[233] reviews seven studies reporting between 31% and 79% confirmation rates, with the lowest figure coming from the most recent study published in 2003. A continuing decrease in the proportion of clinical PID that is confirmed on laparoscopy is probably to be expected in view of the fact that it is now recognised that acute PID is difficult to diagnose because of the wide variation in the symptoms and signs. Many women with PID have subtle or mild symptoms.[63,234] Because of the difficulty of diagnosis and the potential for damage to the reproductive health of women (even by apparently mild or subclinical PID), health-care providers are now advised to maintain a low threshold for the diagnosis of PID.[63,234]

If this is the case, then the most relevant evidence comes from a recently published UK study on a cohort of women with abdominal pain.[235] A total of 112 women were graded as 'almost certain', 'probable' or 'possible' PID, or were 'very unlikely' to have PID on the basis of their symptoms; 42.9% (12/28) of those with grades as having at least probable PID were confirmed on laparoscopy, and 87% of these had high titres of specific IgG antibody to CT. If a grading of 'possible PID' is included, 19 out of 56 (33.9%) were confirmed. The definition of 'probable' here accords with the definitions used in our analyses of routine UK data from HES and GPRD (see *Chapter 6*). Therefore, the rate at which women with 'probable/definite PID' acquire salpingitis $\lambda_{a,1}^{salp}$, $\lambda_{a,2}^{salp}$ can be estimated by applying the $\phi^{salp|PID}$ to the clinical PID incidence rates estimated above. We assign an informative prior to $\phi^{salp|PID}$ based on the data, *Beta*(12,16):

$$\lambda_{a,1}^{salp} = \rho.\lambda_{a}^{ALL\ PID}\phi^{salp|PID}$$
$$\lambda_{a,2}^{salp} = \lambda_{a,1}^{PID}.\eta^{GP} \quad (31)$$

Markov model

Figure 17 shows a Markov model of PID development. It is assumed that women start in state 1 at age 15 years and progress through the model until age 44 years. It is a discrete time model with 1-year cycles, programmed in WBDEV and WinBUGS[161,236] (see *Appendix 11*). Expressions for the transition probabilities are shown in *Table 22*, with the parameters defined as in equations 30 and 31. Note that the model shows progression to first salpingitis and subsequent PID episodes. Exactly the same structure is used to assess numbers of PID episodes, or number of salpingitis as described later in the section.

To generate estimates of the proportion of women who have ever had one, two or three or more PIDs, we assume that further PIDs occur within a 2-year time period. The logic of the model is that individuals start in state 1 (no previous salpingitis), and progress to state 2 (salpingitis 0–1 years) at the constant incidence rate $\lambda_{a,1}^{Salp}$. The probability of a transition to state 2 with a 1-year period is therefore $1-e^{-\lambda_{a,1}^{Salp}}$.

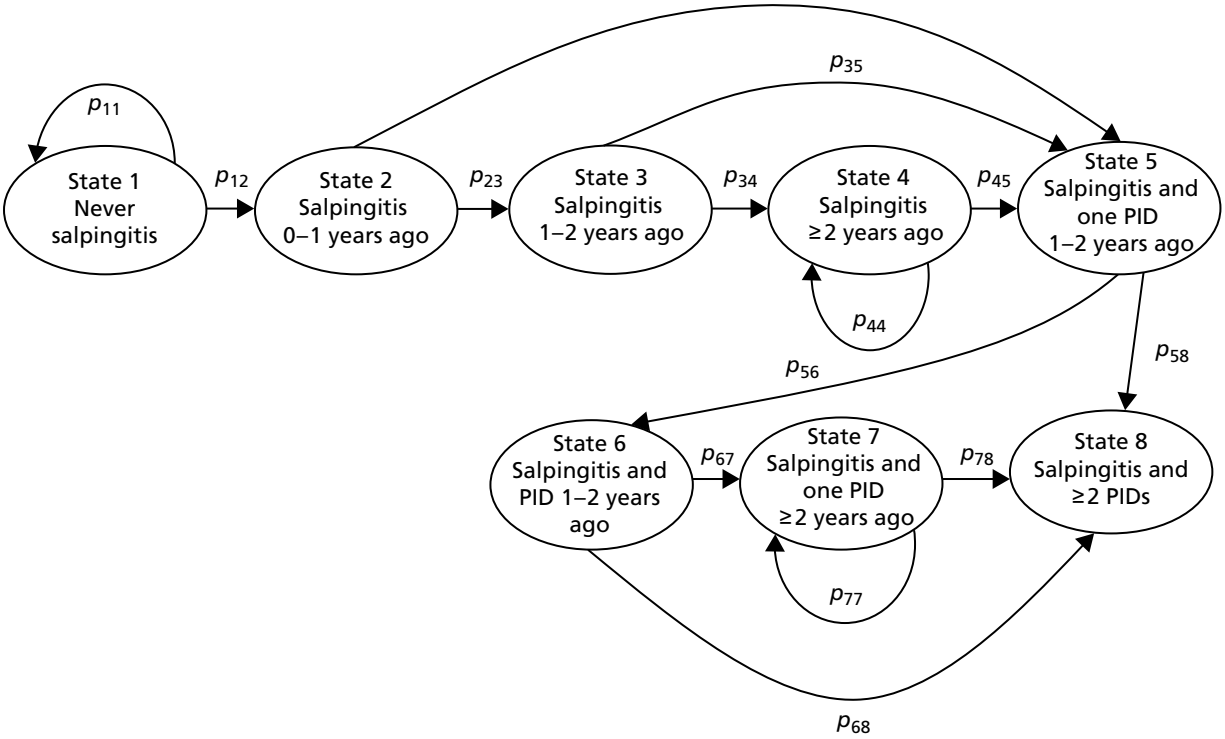

FIGURE 17 Markov model of salpingitis and subsequent PID incidence. Individuals start in state 1 (no previous salpingitis) and progress to state 2 (salpingitis in the last 0–1 year). Once in state 2, women can (a) fail to have a second PID and enter state 3 (salpingitis 1-2 years ago) or (b) have a second PID episode *within a 2-year period* and enter state 5 (salpingitis and 1 PID 0–1 year ago). From state 3 it is also possible to progress to state 5 at the faster 'recurrence' rate or to progress to state 4 (salpingitis 2+ years ago) by failing to experience a further PID episode within 2 years.

TABLE 22 Parameters of the 1-year cycle transition matrix of the Markov model

	Destination states							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
From state								
1	$e^{-\lambda_{a,1}^{Salp}}$	$1-e^{-\lambda_{a,1}^{Salp}}$	0	0	0	0	0	0
2	0	0	$e^{-\lambda_{a,2}^{PID}}$	0	$1-e^{-\lambda_{a,2}^{Salp}}$	0	0	0
3	0	0	0	$e^{-\lambda_{a,2}^{PID}}$	$1-e^{-\lambda_{a,2}^{PID}}$	0	0	0
4	0	0	0	$e^{-\lambda_{a,1}^{PID}}$	$1-e^{-\lambda_{a,1}^{PID}}$	0	0	0
5	0	0	0	0	0	$e^{-\lambda_{a,2}^{PID}}$	0	$1-e^{-\lambda_{a,2}^{PID}}$
6	0	0	0	0	0	0	$e^{-\lambda_{a,1}^{PID}}$	$1-e^{-\lambda_{a,1}^{PID}}$
7	0	0	0	0	0	0	$e^{-\lambda_{a,1}^{PID}}$	$1-e^{-\lambda_{a,1}^{PID}}$
8	0	0	0	0	0	0	0	1

The cell probabilities correspond to the $p_{j',a}$ in equation 33.

Once in state 2, women can (1) fail to have a subsequent PID episode and enter state 3 (salpingitis 1–2 years ago), or (2) have a subsequent PID episode and enter state 5 (salpingitis and 1 PID 0–1 years ago) at a rate $1-e^{-\lambda_{a,2}^{PID}}$. From state 3 it is also possible to progress to state 5 at the faster 'recurrence' rate, or to progress to state 4 (salpingitis 2+ years ago) by failing to experience a PID episode. The subsequent transitions follow the same logic.

Given values of $\lambda_{a,1}^{PID}$, $\lambda_{a,2}^{PID}$ and $\lambda_{a,1}^{Salp}$, the model is run for a cohort of women aged 16–44 years on a 1-year cycle, generating proportions of the female population $\varphi_{k,a}$ in each state k, at each age a. From these state probabilities the proportions of the population who have experienced no salpingitis, or salpingitis with zero, one or two or more subsequent PID episodes can readily be found (see below).

We also validate our results against results from the Lund data set.[33] The study reports data on the distribution of numbers of PID episodes in women who have had salpingitis, for a mean follow-up period of approximately 8 years, for women aged < 25 years, and for women aged ≥ 25 years. The Markov model was run separately for women starting at each of the 22 ages, 16–37 years, in each case starting in state 2. An 8-year time-horizon was used in order to compare results with those observed in the Lund study.[33] The average predicted number of women with a single subsequent PID and women with two or more subsequent PIDs were obtained by averaging across the age ranges 16–24 years and 25–37 years, respectively. Results were obtained for all PID and diagnosed PID only, and these are shown alongside the observed results in the Lund study[33] (*Table 23*).

Model inputs

Besides $\phi^{salpl PID}$ and ρ (see above), the key parameters are:

1. the proportion of all PID that is due to CT, the PEF, for each age group, $\gamma_a^{CT \leftarrow PID}$
2. the incidence of all-cause PID, by age group, $\lambda_a^{ALL\ PID}$
3. the CT re-infection to infection ratio rate for the GP setting, η^{GP}
4. the proportion of PID that is diagnosed in the general population, ψ^{Diag}.

These parameters are all correlated, as they are estimated simultaneously from a common combined data set in *Chapter 7*. After suitable transformations (log transformations for λ_a^{PID} and η^{GP}; a logit transformation for $\gamma_a^{CT \leftarrow PID}$; ψ^{Diag} on the natural scale) the joint uncertainty in these parameters is described by a multivariate normal distribution.

TABLE 23 Distributions of numbers of PIDs from our model after 8 years, posterior mean % in each category (95% CrIs), compared with the findings from the Lund study[33] (see text)

Age	Model		Lund study[33]
	All PID	Diagnosed PID	
16–24 years			
1 PID	66.3 (51.3 to 78.0)	85.4 (78.3 to 90.2)	77.7 (75.1 to 80.3)
2 PID	23.6 (18.3 to 27.2)	12.8 (9.1 to 17.8)	16.0 (13.8 to 18.4)
3+ PID	10.1 (3.75 to 21.5)	1.8 (0.7 to 3.9)	6.2 (4.8 to 7.8)
25–44 years			
1 PID	75.2 (62.3 to 84.7)	90.0 (84.4 to 93.5)	87.0 (82.6 to 90.9)
2 PID	19.3 (13.5 to 25.0)	9.3 (6.2 to 13.6)	11.0 (7.5 to 15.1)
3+ PID	5.5 (1.8 to 12.7)	0.9 (0.3 to 2.0)	2.0 (0.6 to 4.0)

Model outputs

Our primary output from the chapter is $\pi_{m,a}^{S+PID}$, the proportion of women age *a* who have had an episode of salpingitis and zero, one, or two subsequent episodes of PID (indexed by m = 0,1,2 respectively). At age 15 years, all women are assumed to have never had PID so that $\varphi_{1,15} = 1$, $\varphi_{2,15} \cdots \varphi_{8,15} = 0$. If $\varphi_{k,a}$ is the proportion of women at age *a* in state *k*, then:

$$\begin{aligned} 1 - \sum_{m=0,1,2} \pi_{m,a}^{S+PID} &= \varphi_{1,a} \\ \pi_{0,a}^{S+PID} &= \varphi_{2,a} + \varphi_{3,a} + \varphi_{4,a} \\ \pi_{1,a}^{S+PID} &= \varphi_{5,a} + \varphi_{6,a} + \varphi_{7,a} \\ \pi_{2,a}^{S+PID} &= \varphi_{8,a} \end{aligned} \tag{32}$$

where:

$$\varphi_{j,a} = \sum_{j'=1}^{8} \varphi_{j',a-1} \cdot P_{j',j,a} \tag{33}$$

$P_{j',j,a}$ is the transition probability from state *j'* to state *j* where a indexes from ages 16 to 44. The model was run first for all-cause PID, and then separately for diagnosed all-cause PID. Although the true progression through the model is the same, only a proportion ψ^{Diag} of PIDs are observed. The probability of diagnosis for each PID is assumed to be independent of the number of previous PIDs, and it is assumed that no women have more than three PIDs, so the results will be a slight underestimate. As such, the proportion of women observed to have had zero, one, two or three or more PIDs, respectively, $\pi_{n,a}^{obsPID}$, is:

$$\begin{aligned} \pi_{0,a}^{obsPID} &= 1 - (\pi_{1,a}^{obsPID} + \pi_{2,a}^{obsPID} + \pi_{3,a}^{obsPID}) \\ \pi_{1,a}^{obsPID} &= \pi_{1,a}^{PID} \cdot \psi^{Sym} + 2 \cdot \pi_{2,a}^{PID} \cdot \psi^{Sym} \cdot (1 - \psi^{Sym}) + 3 \cdot \pi_{3,a}^{PID} \cdot \psi^{Sym} \cdot (1 - \psi^{Sym})^2 \\ \pi_{2,a}^{obsPID} &= \pi_{2,a}^{PID} \cdot (\psi^{Sym})^2 + 3 \cdot \pi_{3,a}^{PID} \cdot (\psi^{Sym})^3 \cdot (1 - \psi^{Sym}) \\ \pi_{3,a}^{obsPID} &= \pi_{3,a}^{PID} \cdot (\psi^{Sym})^3 \end{aligned} \tag{34}$$

Incidence of all-cause PID and incidence of all-cause salpingitis

Exactly the same model that estimates the proportion of women who have had one, two or three or more episodes of PID can be applied to one, two, or three or more episodes of salpingitis, by substituting all instances of the parameter $\lambda_a^{ALL\ PID}$ with the parameter λ_a^{Salp} as calculated in equation 31 and vice versa.

Incidence of non-CT-related PID and salpingitis

Using our PEFs from the previous chapter, we calculate the incidence rate of non-CT related-PID and non CT-related salpingitis for age *a* as:

$$\lambda_a^{nonCT\ PID} = \lambda_a^{ALL\ PID}(1-\gamma_a^{CT\leftarrow PID})$$
$$\lambda_a^{nonCT,Salp} = \lambda_a^{Salp}(1-\gamma_a^{CT\leftarrow PID}) \quad (35)$$

Exactly the same model as above can then be used to calculate the numbers of women who would ever have had 0, one, two or three or more PIDs, and ever had one, two or three or more diagnosed PIDs, by age, if there were no CT-related PIDs. We make the fundamental assumption that the proportion of salpingitis caused by CT is the same as the proportion of clinical PID caused by CT.

Results

The calibration parameter ρ, the ratio of PID incidence in women who have not had a PID in the last 2 years, and the incidence of all-cause PID, was varied in order to find the value at which model predictions best fitted the age-specific all-cause PID incidence derived from *Chapter 7*. The best fitting value was 0.85 (*Table 24*).

Table 23 shows the correspondence between the results from the Markov model run for an 8-year period and the Lund data.[33] Note that the comparisons between observed and predicted distributions relate most directly to the proportions of PIDs that are second or third (or more) PIDs. The proportion of women who have ever had PID is not relevant to the comparison because only women who have a PID were recruited into the Lund study.[33] Note that there is no reason to expect the credible intervals to agree. The first column shows the proportions of women in the study whom the model predicts would develop one or two or more PIDs, of whatever cause and whether diagnosed or not, during the follow-up period. Column 2 shows how many PIDs would be expected to be observed (diagnosed) in these women. The Lund study[33] results (column 3) lie between the results in the first two columns, which is exactly what is to be expected, as it seems reasonable that subsequent PIDs in women who have had a previous, relatively recent, hospital diagnosed PID are more likely than average to be diagnosed, on the basis that (1) these women will be more likely to recognise the symptoms and (2) such PIDs may be more severe than average. On the other hand, the Lund study[33] is not technically a cohort study: unlike the prospective studies analysed in *Chapter 6* we would not necessarily expect all, or even most, symptomatic PID to be diagnosed.

Table 25a gives the predicted numbers of women who have had zero, one, two or three or more previous PID episodes, whether diagnosed or not, by age. This is shown for both all-cause PID, and for non-CT-related PID. An exactly analogous set of predictions is shown for salpingitis (see *Table 25b*). These tables show, for

TABLE 24 Calibration of the Markov model

Age (years)	From cumulative incidence model	Estimated PID incidence
16–44	1.9 (1.2 to 2.8)	1.8 (1.3 to 2.5)
16–19	2.0 (1.4 to 2.9)	2.1 (1.5 to 2.9)
20–24	3.0 (2.0 to 4.6)	2.8 (2.0 to 3.8)
25–34	2.0 (1.2 to 3.2)	1.9 (1.3 to 2.8)
35–44	1.2 (0.72 to 2.0)	1.3 (0.78 to1.9)

Estimated all-cause PID incidence rate per 100 person-years from *Chapter 7* and predictions of the Markov model with the rescaling factor ρ set to 0.85.

TABLE 25a Predicted age-specific distributions of numbers of episodes: all-cause clinical PID and non-CT-related clinical PID, based on the Markov model

	PID episodes			
Age (years)	0	1	2	3+
All PID				
16–19	95.7 (94.2 to 96.9)	3.89 (2.90 to 5.09)	0.43 (0.19 to 0.81)	0.03 (0.01 to 0.10)
20–24	86.7 (82.5 to 90.2)	10.3 (7.94 to 12.9)	2.38 (1.25 to 3.91)	0.64 (0.16 to 1.69)
25–34	75.3 (67.9 to 81.6)	17.2 (13.7 to 21.0)	5.30 (3.10 to 7.95)	2.16 (0.64 to 5.25)
35–44	66.4 (56.9 to 74.6)	22.2 (18.2 to 26.3)	7.76 (4.78 to 11.15)	7.76 (4.78 to 11.2)
Non-CT related				
16–19	97.9 (95.7 to 100)	1.98 (0.00 to 3.87)	0.13 (0.00 to 0.43)	0.01 (0.00 to 0.03)
20–24	91.6 (86.0 to 97.0)	7.05 (2.79 to 10.9)	1.17 (0.22 to 2.68)	0.22 (0.01 to 0.82)
25–34	82.1 (72.4 to 89.3)	14.2 (9.03 to 19.0)	3.15 (1.38 to 6.43)	1.11 (0.18 to 3.31)
35–44	72.4 (61.3 to 82.4)	19.7 (14.2 to 24.8)	5.73 (2.59 to 9.36)	2.15 (0.45 to 5.85)

TABLE 25b Predicted age-specific distributions of numbers of episodes: all-cause salpingitis and non-CT-related salpingitis, based on the Markov model

	Salpingitis episodes			
Age (years)	0	1	2	3+
All salpingitis				
16–19	98.1 (97.0 to 99.0)	1.85 (1.02 to 2.88)	0.04 (0.01 to 0.14)	0.00 (0.00 to 0.00)
20–24	94.1 (90.7 to 96.8)	5.56 (3.18 to 8.35)	0.34 (0.07 to 0.93)	0.02 (0.00 to 0.10)
25–34	88.6 (82.3 to 93.6)	10.4 (6.11 to 15.0)	0.99 (0.23 to 2.44)	0.10 (0.01 to 0.40)
35–44	83.9 (75.3 to 91.0)	14.2 (8.55 to 20.1)	1.71 (0.45 to 3.96)	0.20 (0.02 to 0.78)
Non-CT related				
16–19	99.1 (97.9 to 100)	0.90 (0.00 to 2.01)	0.01 (0.00 to 0.01)	0.00 (0.00 to 0.00)
20–24	96.3 (93.0 to 98.8)	3.56 (1.16 to 6.50)	0.15 (0.01 to 0.52)	0.01 (0.00 to 0.04)
25–34	91.4 (85.3 to 96.0)	7.92 (3.85 to 12.84)	0.60 (0.10 to 1.69)	0.05 (0.00 to 0.22)
35–44	87.1 (78.6 to 93.5)	11.7 (6.22 to 18.0)	1.15 (0.24 to 3.00)	0.11 (0.00 to 0.48)

Posterior mean% in each category (95% CrIs).

example, that 33.6% of women aged 35–44 years have experienced at least one episode of PID, and 16.1% have experienced at least one episode of salpingitis, again all cause and whether diagnosed or not.

The figures that are taken forward to later chapters (*Table 26*) are the proportions of the population that have experienced at least one episode of salpingitis, followed by zero, one or two or more episodes of PID, these being the risk categories used throughout the Lund study.[33] Here we see that, although 16.1% have experienced at least one episode of salpingitis, 10.7% have experienced one salpingitis episode and no further PID episodes, 3.7% one salpingitis and one further PID episode, and 1.7% one salpingitis episode and two or more further PID episodes.

Figures 18–21 give essentially the same results in 1-year bands from age 16 to 44 years for: PID, undiagnosed PID, salpingitis, and salpingitis followed by zero, one, two or three or more episodes of PID.

TABLE 26 Predicted age-specific distributions of numbers of women with a salpingitis and zero, one or two-plus subsequent PID episodes: all-cause, and non-CT related, based on the Markov model

Age (years)	0 salpingitis	1 salpingitis, 0 further PID episodes	1 salpingitis, 1 further PID episode	1 salpingitis, 2+ further PID episodes
All cause				
16–19	98.1 (97.0 to 99.0)	1.69 (0.93 to 2.64)	0.19 (0.07 to 0.38)	0.01 (0.00 to 0.04)
20–24	94.1 (90.7 to 96.8)	4.59 (2.59 to 6.99)	1.06 (0.46 to 1.96)	0.28 (0.06 to 0.79)
25–34	88.6 (82.3 to 93.6)	8.03 (4.63 to 12.0)	2.44 (1.15 to 4.21)	0.98 (0.25 to 2.55)
35–44	83.9 (75.3 to 91.0)	10.7 (6.32 to 15.7)	3.66 (1.80 to 6.14)	1.69 (0.48 to 4.18)
Non-CT related				
16–19	99.1 (97.9 to 100)	0.86 (0.00 to 1.87)	0.06 (0.00 to 0.20)	0.00 (0.00 to 0.01)
20–24	96.3 (93.0 to 98.8)	3.11 (1.07 to 5.58)	0.51 (0.09 to 1.28)	0.10 (0.00 to 0.37)
25–34	91.4 (85.3 to 96.0)	6.48 (3.24 to 10.4)	1.59 (0.52 to 3.26)	0.50 (0.07 to 1.57)
35–44	87.1 (78.6 to 93.5)	9.31 (5.07 to 14.3)	2.66 (1.01 to 5.08)	0.98 (0.18 to 2.85)

Posterior mean% in each category (95% CrIs).

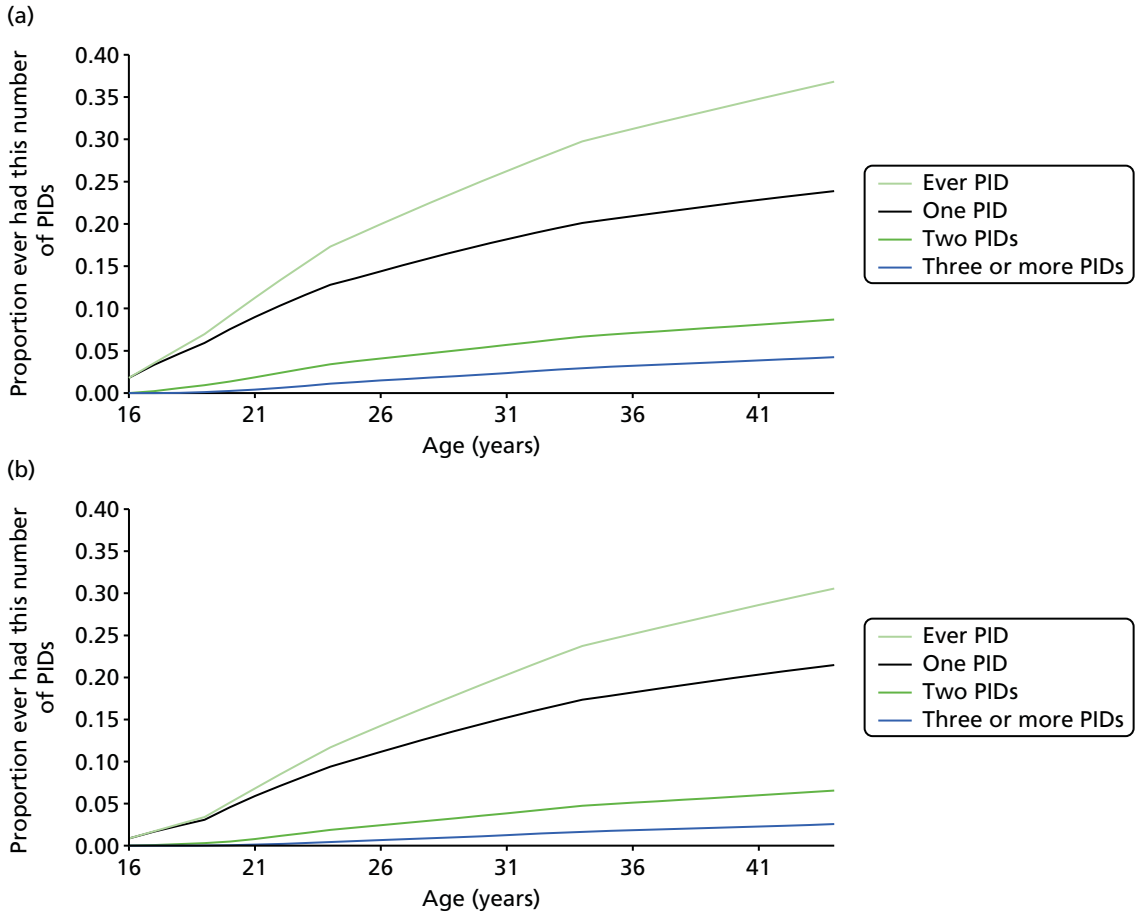

FIGURE 18 Cumulative exposure to PID: proportion experiencing one two and three-plus episodes. (a) All cause; and (b) if no CT.

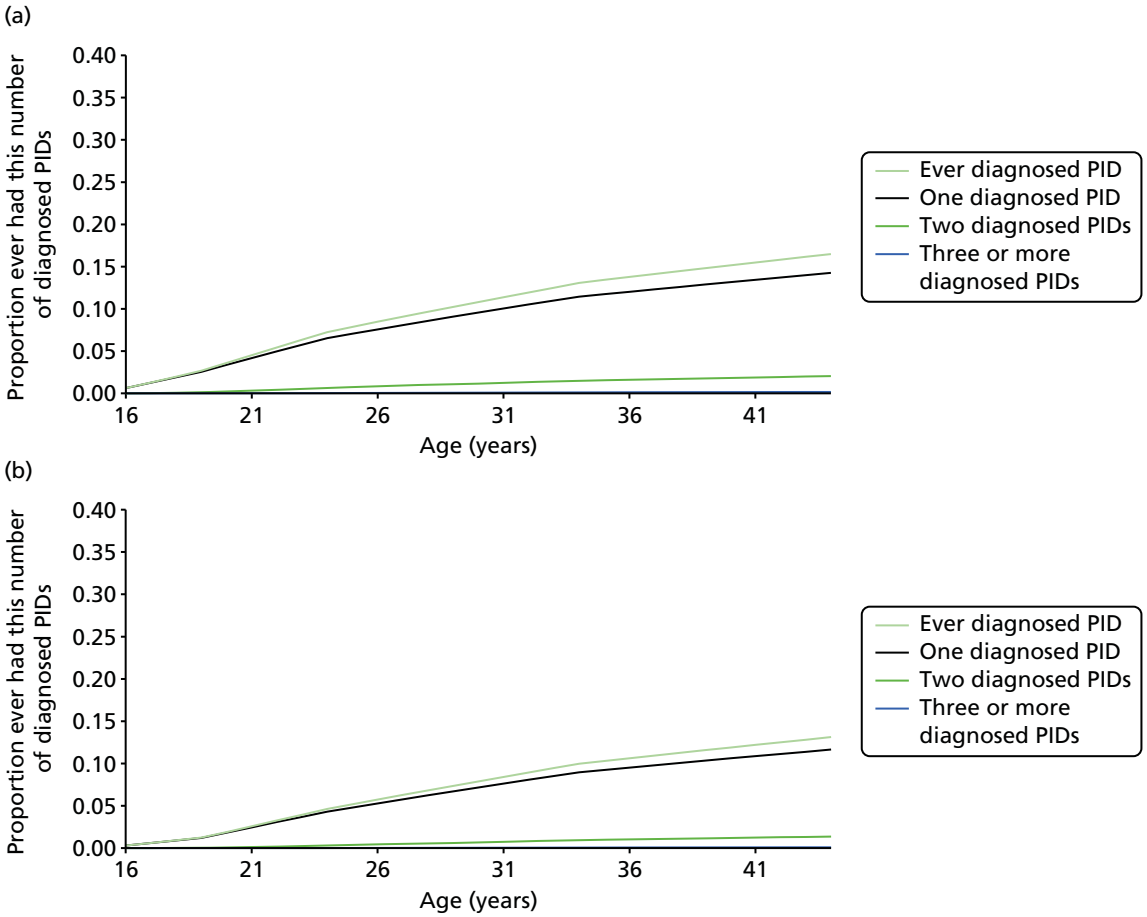

FIGURE 19 Cumulative exposure to diagnosed PID: proportion experiencing one two and three-plus episodes. (a) All cause; and (b) if no CT.

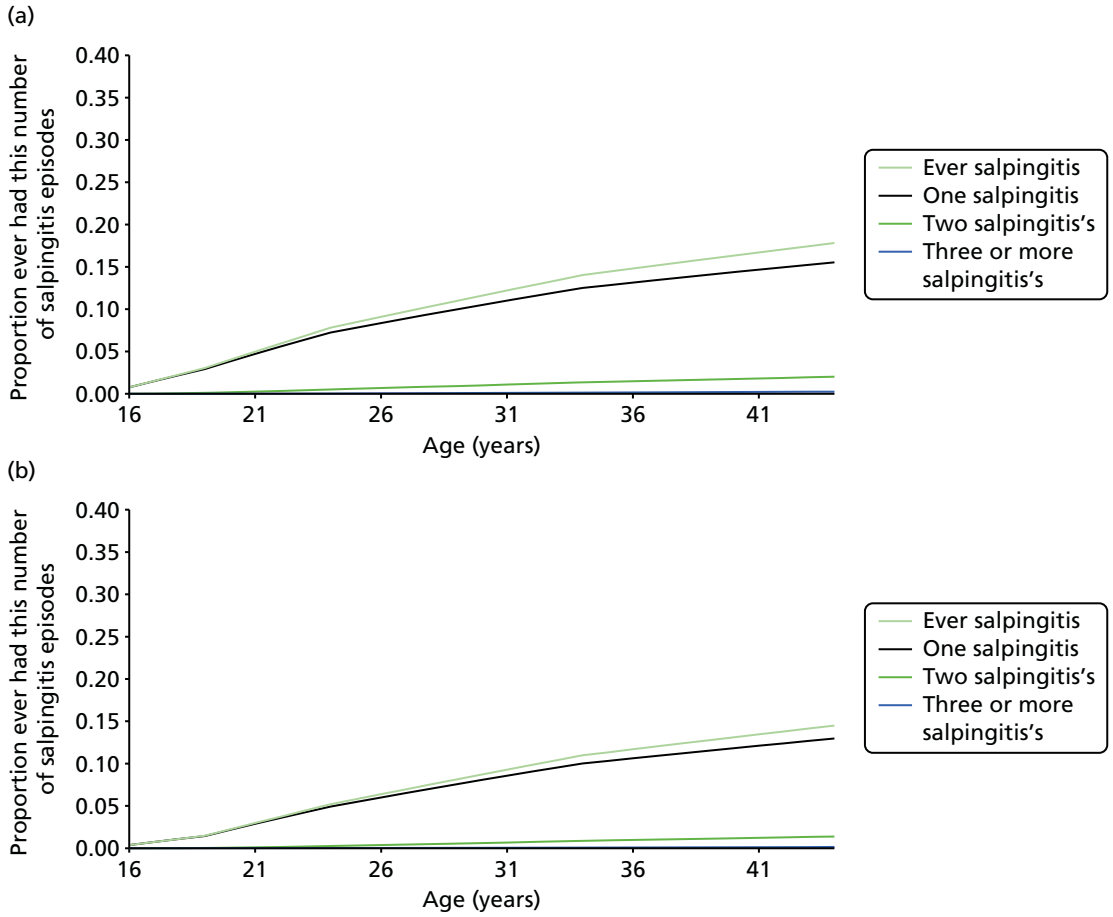

FIGURE 20 Cumulative exposure to salpingitis: proportion experiencing one two and three-plus episodes. (a) All cause; and (b) if no CT.

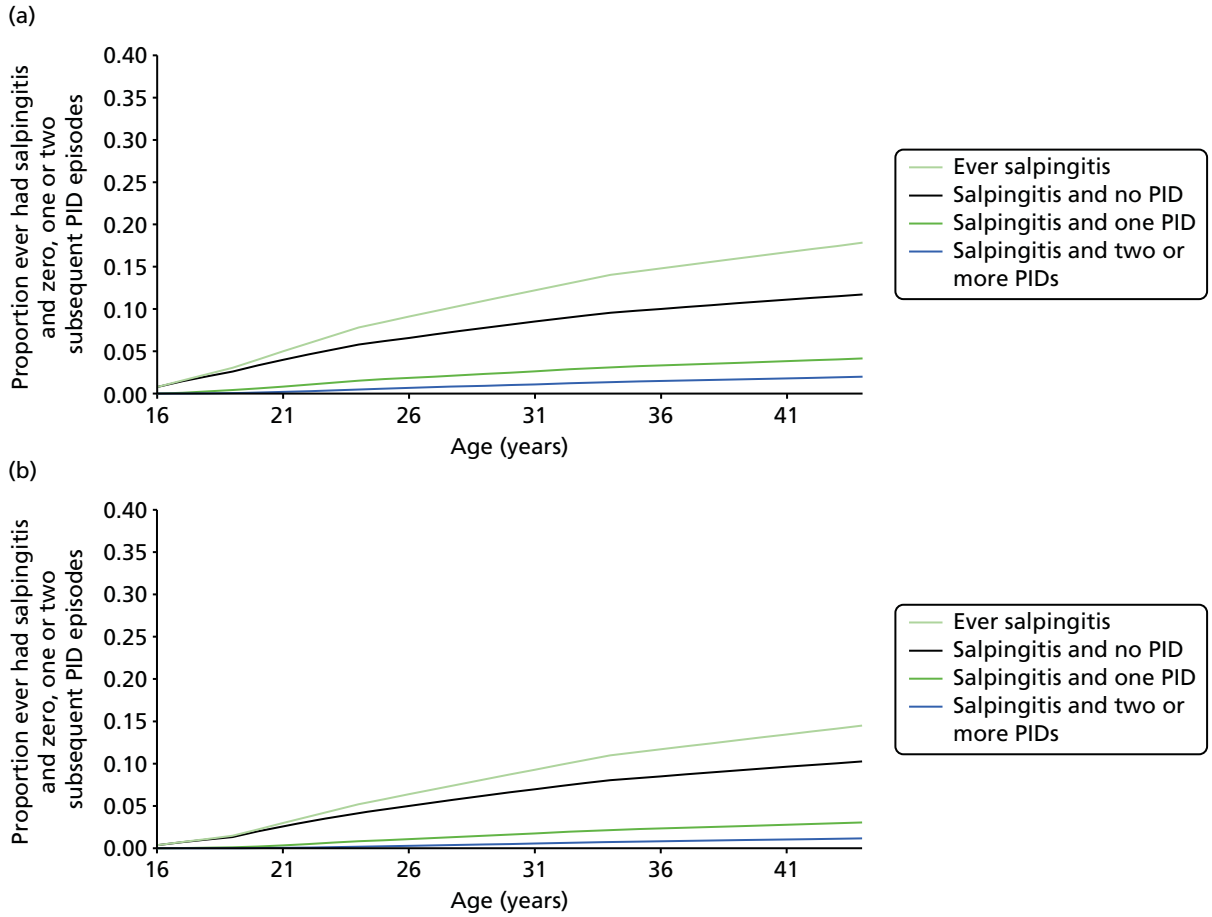

FIGURE 21 Cumulative exposure to salpingitis and subsequent PID episodes: proportion experiencing one episode of salpingitis followed by zero, one and two-plus episodes of PID. (a) All cause; and (b) with no CT.

Discussion

This chapter proposes, as far as we are aware for the first time, a methodology for estimating the proportion of incident PID episodes, and salpingitis episodes, that are first, second or third PIDs, by age. We have applied it to both all-cause PID and salpingitis and non-CT-related PID and salpingitis. The method is based on assumptions about the CT *re*-infection to infection rate ratio, and the length of time after which the re-infection rate applies. With these two assumptions, the results are compared with data on the distribution of second and third episodes in women with an index episode.

In the Lund study that was used for this purpose, 22.1% of women aged 16–24 years with an index salpingitis were observed to have a further PID episode within 8 years, and 6.2% had more than one further episode (see *Table 23*). In women aged 25–44 years, 11% had a further PID episode, and 2% more than one episode. A short-coming of our methodology is that the Lund study follows women whose index PID was sufficiently severe to be not only diagnosed, but also treated in hospital. As a result, our estimate of the proportion of PIDs that are diagnosed in the general population from *Chapter 7* will probably be lower than the proportion in the Lund data, so any comparison must be informal. All we can say is that the Lund results must lie somewhere between our estimates for the numbers of subsequent PIDs and the numbers of subsequent PIDs that are likely to be diagnosed in the general population, which are quite different.

The Lund data on the distribution of repeat PIDs is consistent with data on repeat CT infections in other data linkage studies. For example, in the Uppsala study,[36] of the 3415 CT infections registered in a cohort with a median 15 years' follow-up, 76.9% were first infections. In the nearly 30,000 women in the Norwegian cohort,[215] with an average 8 years' follow-up, 62% CT infections were first infections, 25% second and 13% third or more. Nevertheless, there is no strong empirical basis for choosing the re-infection rate in the GP group from the LaMontagne study,[192] nor for choosing a 2-year period during which the higher re-infection rate applies: both assumptions are somewhat arbitrary. Clearly, other assumptions could be made, which would generate different estimates of the relative proportions of subsequent PIDs that are first, or second-plus episodes. However, the total numbers of PIDs – whether first, second or third – must remain the same as our estimates in *Chapter 7*; similarly, the distribution of the number of subsequent PIDs is required to be consistent with the Lund data. It is therefore to be expected that, given these constraints, similar results will be obtained with different inputs.

An interesting aspect of the modelling approach we have taken is that it correctly captures the age profile of PID, which peaks in the 20- to 24-year-old age band, *not* in the < 20-year-old age band in which CT incidence peaks. As we will see when considering the epidemiology of EP (see *Chapter 9*) and TFI (see *Chapter 10*), the different age profiles of CT and each of its sequelae provide important insights into the aetiology of these conditions, which have perhaps not been adequately addressed in previous work.

We use data from the UK study by Taylor-Robinson[235] to estimate the proportion of PID cases that are salpingitis. Although the study was published fairly recently, the data were collected in the 1990s and patients were diagnosed in hospital, so it is unclear how applicable it is to all PID in 2002. Over time, clinical guidance has changed to treat women with possible or probable PID instead of only treating women with probable PID, so this is likely to be an overestimate for the proportion in all PIDs in 2002. Although we included only PID cases from the GPRD database that are definite and probable, it is unclear whether the Taylor-Robinson study[235] provides an overestimate for clinical PID cases diagnosed in GUM clinics. We also assume the same proportion of salpingitis in undiagnosed PID cases, and there is no real evidence to say whether this is reasonable. On the one hand, undiagnosed women are likely to have less severe symptoms and symptoms are likely to correlate to severity of inflammation and presence of salpingitis. However, laparoscopy identifies the presence of salpingitis at a single point in time. Some of the women in the Taylor-Robison study[235] may have developed inflammation that would be visible on laparoscopy at a later date had they not been treated, as would be the case if they were undiagnosed.

There are some quite severe limitations to our interpretation of the present findings, however. One relates to the fact that our estimates of PID incidence and the age-related population attributable fractions, are based on 2002 data. We are therefore implicitly assuming that age-specific PID incidence is constant over time, although evidence from most countries suggests that it has been declining steadily since the 1970s.[237] Although the numbers of PIDs reported through the HES has remained fairly constant since 2002, the numbers coming through KC-60 returns have increased, and those reported in GPRD have nearly halved.[90]

Given the lack of empirical basis for some of the assumptions made in our analysis of cumulative PID, it is important to compare our quantitative findings with as broad a range of other studies as possible. Our estimates of cumulative incidence of PID seem high, and our estimates of cumulative incidence of diagnosed PID are considerably higher than NATSAL,[195] which reported that 2.2% (1.8% to 2.6%) of female respondents said they have ever been treated for PID compared with our estimate of about 10% in 31-year-olds (see *Figure 18*). However, NATSAL is also highly inconsistent with other UK data sources. The POPI trial[19] observed all-cause PID incidence to be 2% *in a single year*. Furthermore, HES data alone reports a total of approximately 15,000 PIDs in women by the age of 35 years: given that an annual birth cohort would comprise some 300,000 females, unless over half of these episodes occurred in the same women this equates to more than the amount reported by NATSAL.

Recruitment and participation biases in surveys such as NATSAL may selectively under-sample or those who would be considered at increased risk (and some groups at reduced risk). On top of this, there may be a tendency among responders to under-report health problems linked to STD, and it may be that not everyone diagnosed with PID is told this diagnosis and remembers it. The discrepancy between NATSAL and our results is, nevertheless, large and requires further investigation.

Our projections can also be compared with the 2002 US National Survey of Family Growth in which 5.1% of women aged 16–44 years reported having been treated for PID.[238] This figure is sharply down on the 1995 Survey, which reported 8% had been treated for PID, with 11% in the 1988 and 14% in the 1982.[238] Our average estimate for this age range is around 10% (see *Figure 19a*).

The Scandinavian linkage studies provide another opportunity to 'reality check' the predictions. In the Uppsala study, the cumulative incidence of PID was reported as 3.9% by 35 years[36] compared with our estimate of 13.4% for all (diagnosed) PID (see *Figure 19a*). The Uppsala figures, however, cover only hospital-diagnosed PID. It is not known what proportion of PID was diagnosed and treated outside the hospital setting. However, we can obtain a comparable estimate from our model by substituting our age-specific estimates of the proportion of PID cases diagnosed in hospital (see estimation of $\kappa_{1,a}$ in *Chapter 9*) for ψ^{sym} in equation 34. This gives an estimate of 4.6% for the proportion of women aged 35 years who have ever been diagnosed with PID in hospital, which is close to the Uppsala figure.

Summary of assumptions and findings

Summary of assumptions

1. The distribution of zero, one and two or more subsequent PID episodes in women with salpingitis in the UK approximates that found in the Lund study.
2. The probability that PID is diagnosed does not depend on the number of previous episodes, nor on whether previous episodes would have confirmed as salpingitis on laparoscopy.
3. It is assumed that the re-infection rate for CT applies equally to all other causes of PID and salpingitis.

Summary of findings

1. 33.6% of women age 35–44 years have experienced at least one episode of PID (diagnosed or not).
2. 16.1% of women age 35–44 years have experienced at least one episode of salpingitis (diagnosed or not).

Chapter 9 Pelvic inflammatory disease and ectopic pregnancy

Objectives

To:

1. use routinely collected data to estimate the proportion of pregnancies that are ectopic in the UK, by age
2. apply the information on group-specific risks of EP reported in the Lund study (severity of salpingitis, number of diagnosed PID episodes) to estimates of the distribution of these risk factors in the UK, to predict the proportion of salpingitis-related pregnancies that are ectopic in the UK
3. compare predicted and observed proportions of pregnancies that are ectopic
4. estimate the proportion of EP attributable to salpingitis
5. estimate the proportion of EP attributable to CT.

Introduction

This chapter looks at the evidence relating PID, and particularly salpingitis, and EP. As we have seen, the most important findings, from the Lund studies,[33,110,124,232] are that the risks of subsequent EP and TFI depend on salpingitis severity at laparoscopy, and that they increase with number of episodes of PID. However, these observations relate exclusively to women with PID diagnosed and treated in hospital, and were made some 30 years ago in Sweden. We therefore need to address the huge evidence gap regarding the role of PID that is diagnosed and treated outside the hospital setting and the role of undiagnosed PID. Evidence presented in *Chapters 2* and *7* suggested that these make up the majority of PID episodes.[103] We also need to obtain contemporary estimates that are relevant to the UK.

In previous chapters we have been able to set out a mathematical model a priori, based for example on the well-understood relationship between incidence, prevalence and duration (see *Chapters 4* and *5*) or based on relatively familiar Markov processes (see *Chapter 6*). We have then been able to assess whether the different evidence sources were consistent *under a mathematical model* that was not only *prespecified* but whose form and structure could not be influenced by the evidence available. In this chapter and the next we are unable to follow such a rigorous approach because the natural history of PID and the mechanisms that lead from PID to EP and TFI are far less well understood.

We can, however, retain the essential element of our approach, which is the critical examination of a wide range of different types of evidence bearing on the relationship between a small set of outcomes within a coherent mathematical model, in this case CT, PID, salpingitis and EP. However, rather than assessing the consistency of the evidence sources under a prespecified model, we instead try to answer the question: 'under what models can these data sources be considered consistent?' The exercise therefore takes on more of a 'hypothesis generation' role. However, as shown in *Chapter 2*, there is a single common pathway between PID and EP, and PID and TFI. Therefore, we would expect any structural assumptions we make for EP to also hold for TFI. This will provide an opportunity to assess consistency and the compatibility of findings on EP risk and TFI risk in the concluding *Chapter 12*.

We have referred previously to the strong 'dose–response' relationship reported in prospective studies following women with hospital diagnosed PID. The Lund studies,[33,110,124,232] in particular, followed women for approximately 8 years after hospitalisation for PID. The proportion of women whose first subsequent

pregnancy was an EP increased systematically with severity of salpingitis as determined by laparoscopy, and also with the number of diagnosed subsequent PID episodes.[124] A similar pattern of findings was reported for TFI.[33]

A record linkage study in the USA following high-risk women with a diagnosed CT infection over an average 6-year period, again reported a similar increase in risk of EP, but this time in relation to the number of previous CT infections rather than the number of diagnosed PID episodes.[107] These studies constitute overwhelming evidence of the causal role of PID in reproductive damage, and point to the significance of *cumulative* exposure to risk factors, as described in *Chapter 2*. In addition, the similarity of the 'dose–response' effect attaching to CT and PID suggests that the risk of EP and TFI may be approximately the same whatever the cause of the PID, although the agreement could be coincidental if the aetiology changes with age in a way that approximately cancels out any differences.

Review and interpretation of evidence sources

In this section we describe what data sources are potentially available and we develop a rationale that explains which ones we have selected for use. At the same time, we develop a framework that embraces the different types of study: prospective studies of EP risk in women with diagnosed PID, routine data on the incidence of EP in England, and retrospective studies of the PEF of EP that is caused by PID.

Prospective evidence: ectopic pregnancy following pelvic inflammatory disease

We identified studies from a recent systematic review of the sequelae of CT,[98] from the knowledge of collaborators, and as part of a wider search of the literature on CT. We exclude a study by Buchan *et al.*[239] because it does not report data on number of pregnancies (see below). *Table 27* summarises the results of the six prospective studies[33,228,240–243] following women forward from PID to a next pregnancy. These studies have very different follow-up times. However, we can control for this approximately by estimating the proportion of pregnancies that are ectopic, $\varphi^{EP|preg}$. Study-specific crude estimates are presented in column 7. Apart from the small study by Heinonen *et al.*,[240] in which EP rates are unaccountably low, estimates from studies following women with laparoscopically confirmed salpingitis are reasonably homogeneous, at ≈7–10%. The study by Ness *et al.*[228] estimates a much lower proportion (≈1%). It follows women with mild to moderate PID, not laparoscopically confirmed salpingitis, and the most serious 3% of cases are excluded. The results from the prospective studies are otherwise broadly similar.

Our quantitative analysis of the prospective risk of EP following salpingitis is based solely on the Lund study,[124] for four reasons: it is by far the largest study; it is the only study to include a control group; uniquely it presents EP risk by age, severity of salpingitis in women with PID, and numbers of episodes of PIDs; and also it delivers a related set of estimates of the risk of TFI following hospital diagnosed PID with salpingitis. This last feature is especially significant, as it provides an opportunity to check the consistency between observed and predicted EP and TFI frequency in the UK on the same basis.

Application of the Lund results to the epidemiology of pelvic inflammatory disease in the UK

As described in *Chapter 8*, we assume that the proportion of diagnosed 'probable/definite' PID that would be accompanied by salpingitis if examined on laparoscopy is 0.429 (95% CrI 0.27 to 0.63). We further assume that this proportion is the same regardless of PID severity, clinical presentation, place of diagnosis or whether it is diagnosed or not. Note that this implies the existence of undiagnosed but symptomatic salpingitis, and asymptomatic or 'silent' salpingitis, as discussed by Wolner-Hanssen.[103] We assume that only PID that would be confirmed under laparoscopy (salpingitis) can cause EP or TFI. This is supported by observations in the control group in the Lund study[33] (women with hospital-diagnosed PID but no visual evidence of laparoscopy) that 0 out of 601 women developed TFI. The EP rate in this group was slightly higher than the general population but there is considerable uncertainty (only six cases).

TABLE 27 Prospective studies of the risk of EP following PID

First author	Setting and time frame	Mean follow-up time (months)	Arm	No. of pregnancies	No. having an EP	Crude probability a pregnancy is ectopic (95% CrI)	PID diagnostic tool
Westrom[33]	University Hospital, Lund, Sweden (1960–84)	94.3	Case	1100	100	0.09 (0.08 to 0.11)	Laparoscopy
		80.1	Control	439	6	0.01 (0.01 to 0.03)	Laparoscopy
Ness[228]	13 US clinical sites (1996–9)	35	Case	338	5	0.01 (0.01 to 0.03)	Clinical (hospital) presentation [excludes most serious (3%)]
Heinonen[240]	Hospital of Tampere, Finland (1983–8)	125	All cases	56	1	0.02 (0.00 to 0.09)	Laparoscopy
			Mild	29	1	0.03 (0.01 to 0.18)	
			Severe	27	0	0.00 (0 to 0.13)	
Safrin[241]	San Francisco General Hospital (1985)	36–48	Case	44	3	0.07 (0.03 to 0.19)	Clinical (hospital) PID, salpingitis
Gerber[242]	Rostock, Germany (1984–9)	76	All cases	54	5	0.09 (0.04 to 0.20)	Laparoscopy
			Mild	19	0	0.00 (0.00 to 0.18)	
			Moderate	16	3	0.19 (0.07 to 0.46)	
			Severe	19	2	0.11 (0.03 to 0.33)	
Bernsteine[243]	Ohio, USA (not reported)	76	All cases	82	13	0.16 (0.10 to 0.26)	Laparoscopy
			Mild	56	8	0.14 (0.08 to 0.26)	
			Moderate	19	2	0.11 (0.03 to 0.33)	
			Severe	7	3	0.43 (0.18 to 0.81)	

Outline of prospective analysis

The prospective analysis of EP proceeds in the following four steps:

1. A logistic regression model is fitted to the data available from the Lund study (*Table 28*),[124] to estimate the risks that the pregnancy will be ectopic $\eta_{m,s,a}$ in women whose index salpingitis of severity s at age a was followed by $m = 0,1,2+$ subsequent episodes of PID. The non-salpingitis control group informs an age-specific baseline $\theta^{EP|preg}_{Control,a}$, the probability that a pregnancy is ectopic in age group a in women with no previous salpingitis.
2. Because the Lund analysis relates *only* to hospital-diagnosed PID, we explore a range of five scenarios that relate the EP risks $\theta^{EP|Salp,preg}_{m,d,a}$ following $d =$ hospital diagnosed PID (*HD*), PID diagnosed outside of hospital (*nHD*), or undiagnosed PID (*UD*), to the EP risks at different severities s adjusted by the proportion of PID estimated to be salpingitis. These five models can be thought of as 'mappings' between the $\eta_{m,s,a}$ and the $\theta^{EP|Salp,preg}_{m,d,a}$ parameters.
3. The proportion of PID at age a in the categories *HD*, *nHD*, and *UD*, $\kappa_{d,a}$, is derived from the UK routine data on PID and information on the proportion of PID that is diagnosed, ψ^{Diag}. This is multiplied into the proportion of population prevalence of salpingitis in age band a, followed by $m = 0$, 1, 2+ subsequent PID episodes $\pi^{S+PID}_{m,a}$ from *Chapter 8*, to form $\pi^{S+PID}_{m,d,a}$, which has the same interpretation but breaks the population into nine groups by age, diagnosis setting, and number of subsequent PIDs.
4. Estimates of the population proportion of conceptions that are ectopic because of salpingitis are formed by weighting the estimates $\theta^{EP|Salp,preg}_{m,d,a}$, generated by the five 'mapping' models (see below), by the population proportions $\pi^{S+PID}_{m,d,a}$.

The full prospective synthesis is shown in the form of a DAG in *Figure 22*.

TABLE 28 Proportion of pregnancies that are ectopic by number of PIDs, severity and age group: comparison of observed marginal risks with those predicted from the logistic regression model

Risk factor	Original data, Westrom:[124] no. with EP/total pregnancies	Crude estimates as published (95% CI)	Marginal estimates based on regression model (95% CrI)
Control group	6/439	0.014 (0.005 to 0.026)	0.014 (0.005 to 0.026)
No. of PIDs			
One	61/912	0.067 (0.052 to 0.084)	0.067 (0.058 to 0.077)
Two	24/148	0.162 (0.108 to 0.225)	0.163 (0.107 to 0.228)
Three	15/39	0.384 (0.240 to 0.54)	0.382 (0.242 to 0.536)
Age (years)[a]			
16–24	39/713	0.055 (0.039 to 0.073)	0.054 (0.042 to 0.067)
25–44	22/199	0.111 (0.071 to 0.158)	0.113 (0.077 to 0.153)
Severity[b]			
Mild	7/309	0.023 (0.009 to 0.042)	0.021 (0.008 to 0.039)
Moderate	19/420	0.045 (0.028 to 0.067)	0.045 (0.029 to 0.064)
Severe	35/185	0.191 (0.138 to 0.251)	0.195 (0.150 to 0.242)

a Women with one PID only, read from figure 2 in Westrom *et al.*[124]
b Taken from table 2 in Westrom *et al.*[124] first pregnancy only.

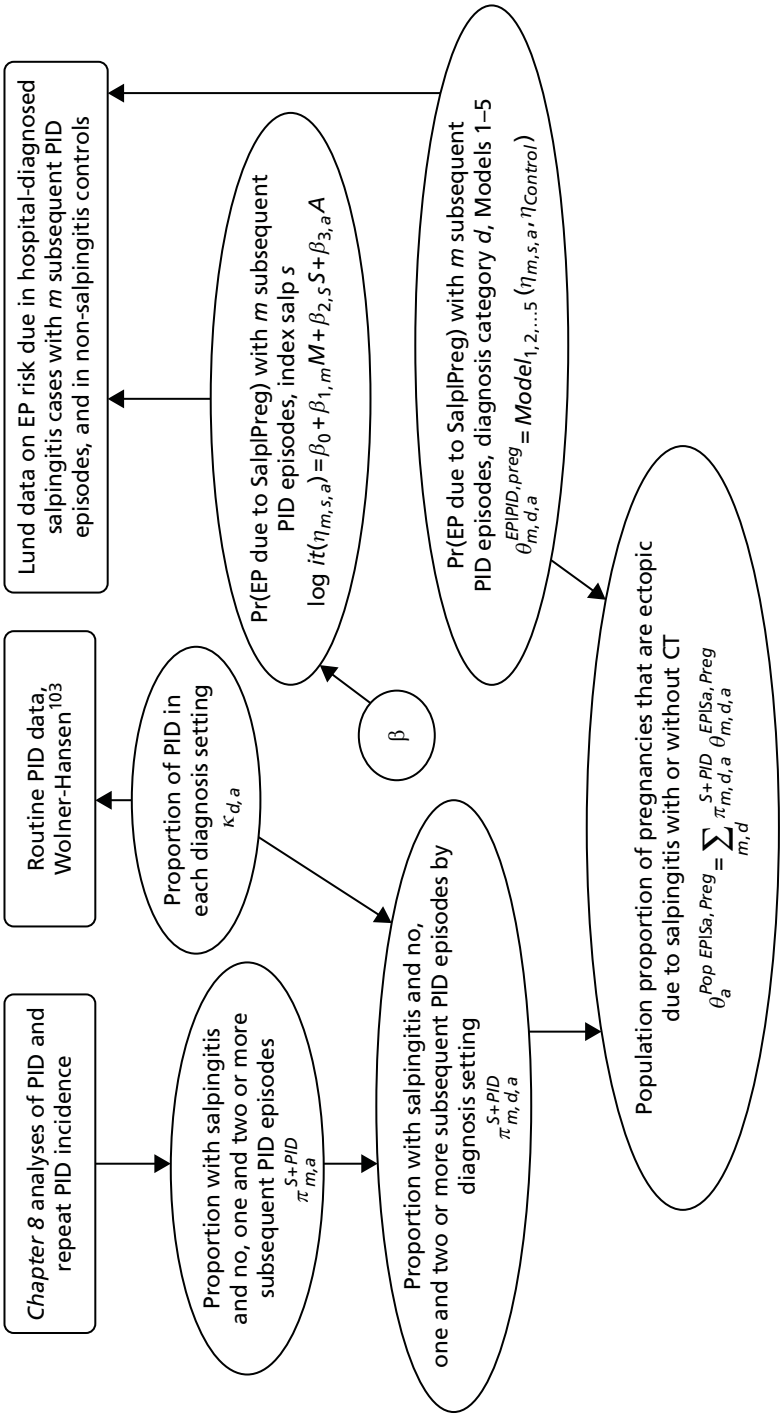

FIGURE 22 Directed acyclic graph of prospective analysis of EP.

Five models relating the Lund estimates in hospital-diagnosed pelvic inflammatory disease to other diagnostic settings

We fit a series of five models with different assumptions about the relation between the risk of ectopic pregnancy reported in the Lund study[124] and the risk in the UK in three groups: (1) hospital diagnosed salpingitis; (2) salpingitis not diagnosed in hospital; and (3) undiagnosed salpingitis. In all five models, we assume that women with hospital diagnosed PID have the same average risk as women (of the same age) in the Lund study. The following five models can be ranked in order from the highest predicted progression rate. To assist distinguishing between the models we use the abbreviations 'Av' for the average progression rate observed in the Lund study, and 'Mi' for the rate in the 'mild' salpingitis.

Model 1 (Av, Av, Av): Lund average for all salpingitis following PID. Here, salpingitis in those with PID diagnosed outside hospital and salpingitis in undiagnosed PID are assumed to have the same risk of EP as hospital diagnosed PID in the Lund study.

Model 2 (Av, Av, Mi): Lund average for salpingitis following diagnosed PID, Lund 'mild' for undiagnosed PID. Salpingitis in all diagnosed PIDs, whether diagnosed in hospital or not, carries the same EP risk as the Lund study; salpingitis in undiagnosed PID has the same level of EP risk as the 'mild' salpingitis group in the Lund study.

Model 3 (Av, Mi, Mi): Lund average for salpingitis following hospital diagnosed PID, Lund 'mild' for salpingitis in GP diagnosed and undiagnosed salpingitis. In this model, salpingitis associated with hospital-diagnosed PID progresses to EP at the average rate observed in the Lund study; salpingitis associated with non-hospital diagnosed and undiagnosed PID progress at the Lund study 'mild' rate.

Model 4 (Av, Mi, Mi/2): Lund average for salpingitis following hospital diagnosed PID, Lund 'mild' for non-hospital-diagnosed. Salpingitis in undiagnosed PID progresses at a rate that is half-way between the 'mild' rate and zero.

Model 5 (Av, Mi, 0): Lund average for salpingitis following hospital diagnosed PID, Lund 'mild' for non-hospital-diagnosed, no risk following salpingitis in those whose PID is undiagnosed.

This choice of models was based on a presumption that the risk of EP will be lower in PID that is diagnosed outside of hospital, and the same or lower again in undiagnosed PID. This is based on the assumption that the risk of poor reproductive outcomes will be lower in women with less severe clinical symptoms, who will be less likely to be diagnosed. This assumption is not necessarily correct. It is also perfectly possible that the milder the clinical manifestation of PID, the later the diagnosis, and the more damage that follows.

In addition to severity of salpingitis, the numbers of subsequent PID episodes also impact on EP risk. Below we describe how we derive predictions for the EP risk attaching to each of the 18 ($3 \times 3 \times 2$) number of subsequent PID episodes, severity and age strata from the information that has been published.

Distribution of pelvic inflammatory disease diagnostic status in the English population

In order to apply the prospective risks informed by the Lund studies to predict EP outcomes in England, it is necessary to have estimates of the risk factor distributions, by age. Estimates of the proportion of women in age band *a* who have had an episode of salpingitis that has been followed by $m = 0,1,2+$ episodes of PID, $\pi_{m,a}^{S+PID}$, where obtained in *Chapter 8*. However, we need to consider the proportions $\kappa_{d,a}$ of women in age band *a* who were hospital diagnosed ($d = HD$), diagnosed elsewhere ($d = nHD$) or undiagnosed ($d = UD$). To do this, we make several fundamental simplifying assumptions, in the absence of clear information to the contrary, and in order to make the analytic task tractable. First, that the probability a PID is diagnosed in hospital is independent of the number of previous PIDs a woman has had; second that the increase in odds of EP given one or two or more as compared to zero, subsequent PID episodes conditional on having an initial PID that is confirmed as salpingitis is independent of whether or where the initial PID was diagnosed, and clinical severity of the initial salpingitis.

Routine ectopic pregnancy data

The premise behind the use of routine data on EP is that it should be possible to compare the observed risk of EP in pregnant woman of age *a*, based on routine HES data for 2002 (*Table 29*), taking account that not all EP is caused by salpingitis, to what would be predicted from the Lund study.

In order to avoid having to explicitly model the complex processes underlying conception including the competing risk of TFI, we have taken as our outcome the proportion of conceptions that are EPs. This is estimated by dividing the number of EPs in HES (see *Table 29*), by the number of conceptions in England and Wales (*Table 30*), to obtain the proportion of conceptions that are ectopic (*Table 31*).

TABLE 29 Hospital episode statistics EP numbers reported, by age and year[117]

	Year									
Age (years)	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
15	8	10	6	18	9	10	11	20	14	16
16–19	305	321	291	356	361	362	416	411	386	381
20–24	1156	1098	1154	1284	1342	1381	1397	1511	1546	1530
25–34	4471	4317	4340	4283	4461	4520	4674	4864	4768	4895
35–44	1750	1862	1840	1957	2038	2048	2141	2223	2221	2213

TABLE 30 Numbers of conceptions, by age and year[118]

Age (years)	Year									
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
13–15	8100	7900	7900	8000	7600	7900	7800	8200	7600	7200
15–17	41,300	41,000	42,000	42,200	42,200	42,300	41,800	43,000	41,400	38,300
15–19	97,700	96,000	97,100	98,600	101,300	102,300	103,100	106,300	103,300	97,900
20–24	159,000	161,600	167,800	175,300	181,300	185,500	191,200	198,700	198,500	199,500
25–29	209,300	199,300	199,400	199,800	205,100	211,300	222,200	234,800	237,800	242,200
30–34	195,300	196,700	204,300	209,000	209,600	209,200	212,400	211,500	207,100	213,300
35–39	88,700	92,200	98,900	103,100	106,800	110,000	115,400	118,000	115,600	116,500
40+	17,000	17,800	19,600	20,900	22,800	23,600	25,500	26,500	26,500	26,800

TABLE 31 Percentage of conceptions that are ectopic, by age and year

	Year									
Age (years)	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
16–19	0.34	0.36	0.33	0.39	0.39	0.38	0.44	0.42	0.40	0.42
20–24	0.73	0.68	0.69	0.73	0.74	0.74	0.73	0.76	0.78	0.77
25–34	1.11	1.09	1.08	1.05	1.08	1.07	1.08	1.09	1.07	1.07
35–44	1.66	1.69	1.55	1.58	1.57	1.53	1.52	1.54	1.56	1.54

Note
We assume that all conceptions in females aged 13–15 years occur at age 15 years, and all conceptions in females aged 40+ years occur between the ages of 40 and 44.

This conditions out all of the complexities of conception rates, such as their dependency on fecundity and time taken to conceive, and it links directly to the Lund study, in which outcomes are also reported as the proportion of conceptions that are EPs.

Retrospective case–control studies of ectopic pregnancy and pelvic inflammatory disease/salpingitis

We searched for case–control studies comparing history of salpingitis in cases with EP and pregnant controls with no EP. A meta-analysis of studies on risk factors for EP[147] identifies previous gonorrhoea, serological evidence of previous CT, PID, smoking, number of previous sexual partners, and previous EP and tubal surgery as risk factors for EP. Another risk factor that was not included in this study, but which has received considerable attention, is previous use of an IUD.[244,245] There was, however, considerable heterogeneity between studies, with different risk factors being reported in different studies. Therefore, as in our search for case–control studies looking at PID, we limited attention to studies conducted in Europe during the 1990s or later in order to exclude studies for which the spectrum of causative agents for EP, such as smoking, were likely to have been different. We also only consider studies in which PID history was assessed through laparoscopy because of the poor specificity of clinically diagnosed PID as a proxy for salpingitis. Papers were identified through the knowledge of collaborators and our wider review of the literature, from modelling studies and economic evaluations, and citation and reference searches of a systematic review,[147] and any studies that were identified as relevant.[246,247] Retrospective studies were not considered in the recent series of systematic reviews by the CDC. We identified two studies,[246,247] both carried out in France.

The study by Bouyer *et al.*[246] is the largest and most recent. It provides estimates of both the crude OR for probable PID (fever, abdominal pain, vaginal discharge) and for laparoscopically confirmed PID (salpingitis). Estimates were based on a logistic regression model that adjusts for a number of possible confounders. The unadjusted OR (5.42, 95% CI 2.60 to 11.28) is likely to be an overestimate, in view of the likely positive correlation between risk factors, the adjusted estimate (3.39, 95% CI 1.44 to 7.96) is likely to be too low as a result of 'over-controlling'. For example, history of infertility is included as a covariate. But, it seems unlikely that a history of infertility causes EP; instead, both of these outcomes will share some of the same risk factors, including PID. We therefore assume that the crude OR represents an upper bound, whereas the adjusted OR from the regression model is a lower bound for the OR for PID. The above figures are for 'confirmed PID' (see below), which constituted 90% of all PID reported in the study. The fact that the ORs for confirmed are higher than for probable PID suggests that the more stringent criteria are better predictors of subsequent EP and therefore a better reflection of the causal process.

The Coste study[247] reports crude ORs of 5.8 (95% CI 2.4 to 14.3) for laparoscopically confirmed PID, and 3.2 (95% CI 1.3 to 7.8) for probable PID. This study[247] also reports an adjusted estimate based on a regression controlling only for age, hospital, current smoking, previous use of IUD, and current contraception method. This adjusted OR (4.7, 95% CI 1.7 to 13.6) seems less likely to be 'over-controlled', and is between the higher and lower estimates from the Bouyer *et al.* study.[246] Rather than attempt a formal synthesis of these estimates we use the higher and lower limits from Bouyer *et al.*[246] in the analyses reported below.

It should be noted that the definition of 'confirmed' PID in both studies is not entirely clear. It appears that confirmation could be either by laparoscopy or positive CT serology. The circumstances under which laparoscopy was routinely carried out in cases of suspected PID is not specified, nor is it known what proportion of 'confirmed PID' was confirmed by serology alone or what cut-offs and tests were used.

As EP is a rare disease, we may use the adjusted OR from these studies to estimate the proportion of EP attributable to salpingitis, based on the standard formula (see *Chapter 3*). Values for the prevalence of exposure to salpingitis are derived from *Chapter 8*.

Statistical methods

Risk of ectopic pregnancy in pregnant women as a function of age, PID severity and number of PID episodes, based on Lund data

We fit the following logistic regression model to the data in *Table 28* on EP outcomes in patients with an index salpingitis:

$$logit(\eta_{m,s,a}) = \beta_0 + \beta_{1,m}M + \beta_{2,s}S + \beta_{3,a}A \tag{36}$$

where $\eta_{m,s,a}$ is the risk factor specific risk of EP, β_0 is the (baseline) log odds that a pregnancy will be ectopic following a mild salpingitis episode at age < 29 years and not followed by any further PID episodes. Similarly, $\beta_{1,m}$ is the log odds ratio (LOR) of EP for a woman with m subsequent PID episodes compared with none, $\beta_{2,s}$ is the LOR for a woman whose index salpingitis was moderate, or severe compared with mild, and $\beta_{3,a}$ is the log OR for age group a compared with age group 1. The dummy variables M,S, and A denote M = one or two further episodes, S moderate or severe, A = age ≥ 29 years.

Note that the risks of EP due to salpingitis are considered to be *additional* to a baseline risk $\eta_{Control}^{EP|preg}$ experienced by the general population, which we estimate from the non-salpingitis PID controls in *Table 28*.

The parameters of equation 37 model can be identified from the published data[124] that are available (see *Table 28*) but we are unable to identify interaction terms. Although this does not necessarily oblige us to assume there are no interactions, in the absence of any evidence, this is the assumption that is made. Finally, we note that, in the Lund study,[124] the results are reported by the age groups of < 25 years and ≥ 25 years. This refers to age-at-diagnosis, at study entry, and our analysis requires knowledge of the age when EP occurs. The study has almost a 9-year follow-up, so we assume that EP happens on average just over halfway through and that the ≤ 24 age group represents women who have EP by the age of 29 years.

We assume that the numerators, the number of EPs, in the Lund data (see *Table 28*), by number of PIDs $r_{1,m}$, by severity $r_{2,s}$, and by age $r_{3,a}$, have binomial likelihoods:

$$r_{1,m} \sim Binomial\ (p_{1,m}, n_{1,m})$$
$$r_{2,s} \sim Binomial\ (p_{2,s}, n_{2,s})$$
$$r_{3,a} \sim Binomial\ (p_{3,a}, n_{3,a})$$

where:

$$p_{1,m} = \frac{\sum_{a=1}^{2}\left(\frac{\sum_{s=1}^{3}\eta_{m,s,a}\cdot n_{3,a}}{\sum_{s=1}^{3}n_{2,s}}\right)}{\sum_{a=1}^{2}n_{3,a}}, \quad p_{2,s} = \frac{\sum_{a=1}^{2}\eta_{1,s,a}\cdot n_{3,a}}{\sum_{a=1}^{2}n_{3,a}}, \quad p_{3,a} = \frac{\sum_{s=1}^{3}\eta_{1,s,a}\cdot n_{2,s}}{\sum_{s=1}^{3}n_{2,s}} \tag{37}$$

The assumption of independence of the likelihoods is not quite correct, as the same data are being used in different breakdowns, and this will have the effect of exaggerating somewhat the precision of our estimated risk coefficients.

The parameters β_0, $\beta_{1,m}$, $\beta_{2,s}$ and $\beta_{3,a}$ are given flat normal priors.

Distribution of number of salpingitis episodes and number of subsequent PID episodes and severity in the English population

We embed the WBDEV code described in *Chapter 8* directly into the program to estimate $\pi_{m,a}^{S+PID}$ from parameters relating to PID incidence from previous chapters. These are $\lambda_{a,1}^{PID}$, the incidence of a first PID of any cause, diagnosed or not; η^{GP} the ratio of CT re-infection to infection rates in the GP group; ψ^{Diag} the proportion of PID that is diagnosed; and $\phi^{salp|PID}$ the probability of salpingitis given PID. These are given a multivariate normal prior, as described in *Chapter 8*. In parallel, we use the PEF(CT←PID) to estimate the equivalent distribution $\pi_{m,a}^{NoCT,S+PID}$ in a world with no CT.

To estimate $\kappa_{HD,a}$, the age-specific proportion who have experienced *m* subsequent PID episodes following a hospital-diagnosed PID confirmed as salpingitis, we consider the ratio of HES PID episodes by age to the total number of PIDs diagnosed using the same model, as in *Chapter 7*, for the same age groups; this is multiplied by our estimate of the proportion of PID episodes that are diagnosed, ψ^{Diag}, from *Chapter 7*. The proportion diagnosed but not hospital referred $\kappa_{nHD,a}$ can be calculated from this and the proportion diagnosed, which leads directly to an estimate of the proportion undiagnosed $\kappa_{UD,a}$. The calculations are shown below.

Routine data on PID diagnoses in different settings in the general population have binomial likelihoods:

$$\begin{aligned} r_a^{HESPID} &\sim Binomial(p_a^{HESPID}, N_a) \\ r_a^{GPRD} &\sim Binomial(p_a^{GPRD}, N_a) \\ r_a^{kc-60.08} &\sim Binomial(p_a^{kc-60.08}, N_a) \end{aligned} \tag{38}$$

where N_a is the total number of women in age group *a* in the general population in 2002 estimated from census data. KC-60 returns for PID diagnosed in GUM clinics are not available by age before 2008. Data for 2008 are available from Genitourinary Medicine Clinic Activity Dataset (GUMCAD)[248] for approximately 70% of cases. We assume that these cases are a representative sample and that the age distribution within this setting was the same in 2002. The number of cases for each age from GUMCAD are then rescaled to the total number of KC-60 cases in 2002 so:

$$p_a^{kc-60:02} = \frac{13421.p_a^{kc-60:08}}{12117} \tag{39}$$

Using the same model as in *Chapter 7*:

$$\begin{aligned} p_a^{\min} &= p_a^{kc-60} + \max(p_a^{HESPID}, p_a^{GPRD}) \\ p_a^{\max} &= p_a^{kc-60} + p_a^{HESPID} + p_a^{GPRD} \\ X_a &\sim Uniform(p_a^{\min}, p_a^{\max}) \end{aligned} \tag{40}$$

The parameters p_a^{HESPID}, p_a^{GPRD}, p_a^{kc-60} are given *Beta*(1,1) priors and:

$$\begin{aligned} \kappa_{HD,a} &= \psi^{Diag} \frac{p_a^{HES\ PID}}{X_a} \\ \kappa_{nHD,a} &= 1 - \kappa_{HD,a} - \kappa_{UD,a} \\ \kappa_{UD,a} &= 1 - \psi^{Diag} \end{aligned} \tag{41}$$

Finally, we form the full breakdown of population weights, $\pi_{m,d,a}^{S+PID}$, as follows:

$$\pi_{m,d,a}^{S+PID} = \pi_{m,a}^{S+PID} \kappa_{d,a} \tag{42}$$

Prospective estimate of the probability that a pregnancy was ectopic due to salpingitis

An estimate of the proportion of pregnancies that are ectopic as a result of salpingitis in the UK in 2002 from prospective evidence from the Lund study,[124] and estimates of the distribution of numbers and severity of PID, can be generated as follows. First, the proportion of pregnancies that are ectopic because of PID with salpingitis – in women who have had an episode of PID that would confirm as salpingitis on laparoscopy – is constructed by subtracting the proportion of EP pregnancies in women from the same population, the control group *with no index salpingitis*, from a weighted average of the risks in women with mild, moderate and severe PID:

$$\theta_{m,HD,a}^{EP|Salp,preg} = \frac{\sum_{s=1}^{3} \eta_{m,s,a} \cdot n_s^{Lund}}{\sum_{s=1}^{3} n_s^{Lund}} - \theta_{Control,a}^{EP|preg} \tag{43}$$

The weights n_s^{Lund} are the number of women with salpingitis of severity level.

Then the population proportion of pregnancies that are ectopic due to salpingitis, as a function of age, severity of salpingitis, and number of subsequent PID episodes, is constructed as a weighted average using the estimated proportions $\pi_{m,d,a}^{S+PID}$ of women in the UK exposed to $d =$ Hospital Diagnosed (*HD*), Not Hospital Diagnosed (*nHD*), and Undiagnosed PID (*UD*), and the prevalence of zero, one and two-plus PID episodes subsequent to the index salpingitis:

$$\theta_a^{POP,EP|Salp,preg} = \sum_{d=HD,nHD,UD} \sum_{m=0,1,2+} \pi_{m,d,a}^{S+PID} \, \theta_{m,d,a}^{EP|Salp,preg} \tag{44}$$

The five different models of how the EP rates in the Lund hospital-diagnosed PID might apply to GP diagnosed and undiagnosed PID can then be set out in the following way, where 'Av' refers to the average rate reported in the Lund study[124] and 'Mi' to the rate in Mild PID:

Model 1 (Av, Av, Av): Lund average for all salpingitis

$$\theta_{m,nHD,a}^{EP|Salp,preg} = \theta_{m,UD,a}^{EP|Salp,preg} = \frac{\sum_{s=1}^{3} \eta_{m,s,a} \cdot n_s^{Lund}}{\sum_{s=1}^{3} n_s^{Lund}} - \theta_{Control,a}^{EP|preg} \tag{45}$$

Model 2 (Av, Av, Mi): Lund average for salpingitis following all diagnosed PID, Lund 'mild' for Undiagnosed PID

$$\theta_{m,nHD,a}^{EP|Salp,preg} = \frac{\sum_{s=1}^{3} \eta_{m,s,a} \cdot n_s^{Lund}}{\sum_{s=1}^{3} n_s^{Lund}} - \theta_{Control,a}^{EP|preg} \tag{46}$$

$$\theta_{m,UD,a}^{EP|Salp,preg} = \eta_{m,1,a} - \theta_{Control,a}^{EP|preg}$$

Model 3 (Av, Mi, Mi): Lund average for salpingitis following HD PID, Lund 'mild' for salpingitis following nHD diagnosed and UD PID.

$$\theta_{m,nHD,a}^{EP|Salp,preg} = \theta_{m,UD,a}^{EP|Salp,preg} = \eta_{m,1,a} - \theta_{Control,a}^{EP|preg} \tag{47}$$

Model 4 (Av, Mi, Mi/2): Lund average for salpingitis following HD PID, Lund 'mild' for nHD. Salpingitis following UD PID progresses at a rate that is half-way between the 'mild' rate and zero.

$$\begin{aligned} \theta_{m,nHD,a}^{EP|Salp,preg} &= \eta_{m,1,a} - \theta_{Control,a}^{EP|preg} \\ \theta_{m,UD,a}^{EP|Salp,preg} &\sim Uniform\,(0, A_{m,a}) \\ A_{m,a} &= \eta_{m,1,a} - \theta_{Control,a}^{EP|preg} \end{aligned} \tag{48}$$

Model 5 (Av, Mi, 0): Lund average for salpingitis following HD PID, Lund 'mild' for nHD, no risk with salpingitis following UD PID.

$$\begin{aligned} \theta_{m,nHD,a}^{EP|Salp,preg} &= \eta_{m,1,a} - \theta_{Control,a}^{EP|preg} \\ \theta_{m,UD,a}^{EP|Salp,preg} &= 0 \end{aligned} \tag{49}$$

The risk in the control arm $\eta_{Control}^{EP|preg}$ is not reported by age. While keeping the average risk the same as in the controls in the Lund study,[124] we assume that the age distribution of EP in the control group is the same as in the general population in England in 2002.

$$\theta_{Control,a}^{EP|preg} = \eta_{Control}^{EP|preg} \frac{\phi_a^{EP|preg}}{\sum_{a=1}^{4} \phi_a^{EP|preg}} \tag{50}$$

Model for routine ectopic pregnancy data

Data on numbers of conceptions, and hospital episodes of EP for 2002 by age, have a binomial likelihood with the denominator taken from census population estimates for that year.

$$\begin{aligned} r_a^{HES,EP} &\sim Binomial(p_a^{HES,EP}, N_a) \\ r_a^{CONC} &\sim Binomial(p_a^{CONC}, N_a) \end{aligned} \tag{51}$$

The parameters p_a^{HESPID} and p_a^{GPRD} are given effectively uninformative *Beta*(1,1) priors. Finally, the proportion of pregnancies that were ectopic in the English population in 2002 $\phi_a^{EP|preg}$ equals:

$$\phi_a^{EP|preg} = \frac{p_a^{HES,EP}}{p_a^{CONC}} \tag{52}$$

We use the routine data on EP in two ways. The first involves informally comparing the all-cause probability that a pregnancy is ectopic to the predicted proportions due to PID from our model. Second, we combine it with retrospective data to estimate the proportion of EPs caused by PID.

A DAG for the prospective analysis is shown in *Figure 22*, and WinBUGS code is shown in *Appendix 12*.

Model for retrospective data

Lower and upper bounds for the odds of past PID in women with EP compared with women who gave birth (controls) are given informative log-normal priors based on Bouyer *et al.*[246] respectively:

$$\begin{aligned} &\hat{OR_l} = 3.39,\ (95\%CI: 1.44\text{–}7.96) => \log(OR_l) \sim Normal(1.22, 0.19) \\ &\hat{OR_u} = 5.42,\ (95\%CI: 2.60\text{–}11.28) => \log(OR_u) \sim Normal(1.69, 0.14) \end{aligned} \tag{53}$$

The age-specific prevalence of a history of salpingitis is:

$$\pi_{*,a}^{S+PID} = \sum_{m=0,1,2+} \pi_{m,a}^{S+PID} \tag{54}$$

which is calculated using the method described in *Chapter 8* and uses the same MCMC samples as those used in the prospective analysis described above. An analogous estimate can be formed for the prevalence of a history of non-CT-related salpingitis:

$$\pi_{*,a}^{NoCT,S+PID} = \sum_{m=0,1,2+} \pi_{m,a}^{NoCT,S+PID} \tag{55}$$

Estimates of the PEF(Salpingitis←EP) and PEF(Non-CT-related Salpingitis←EP) can then be found using equation 1:

$$\begin{aligned} \gamma_a^{Salp\leftarrow EP} &= \frac{\pi_{*,a}^{S+PID}(OR-1)}{\pi_{*,a}^{S+PID}(OR-1)+1} \\ \gamma_a^{NoCT.Salp\leftarrow EP} &= \frac{\pi_{*,a}^{NoCT.S+PID}(OR-1)}{\pi_{*,a}^{NoCT.S+PID}(OR-1)+1} \end{aligned} \tag{56}$$

Finally, we can form an estimate for the PEF(CT-related Salpingitis←EP). Bearing in mind that an EP can be caused by CT only through a CT-related salpingitis:

$$\gamma_a^{CT\leftarrow EP} = \gamma_a^{Salp\leftarrow EP} - \gamma_a^{NoCT.Salp\leftarrow EP} \tag{57}$$

Lower and upper bounds for all these PEFs based on retrospective data are then calculated by substituting the lower and upper estimates of the OR into the formula. Note, however, this final calculation generates an *under*estimate, because it assumes the distribution of PIDs in women with a salpingitis would be the same in a world where CT does not exist. However, we might expect the shape of the whole distribution to change such that a higher proportion of women who have ever had salpingitis would have fewer subsequent PID episodes. Because of the multiplicative dose–response relationship observed in the Lund study[124] we would expect to observe a lower OR in this population. This would likely be true of the removal of any risk factor. As such the true PEF could perhaps be up to 50% higher than our estimate.

Figure 23 shows a DAG for the retrospective analyses, and WinBUGS code is given in *Appendix 12*.

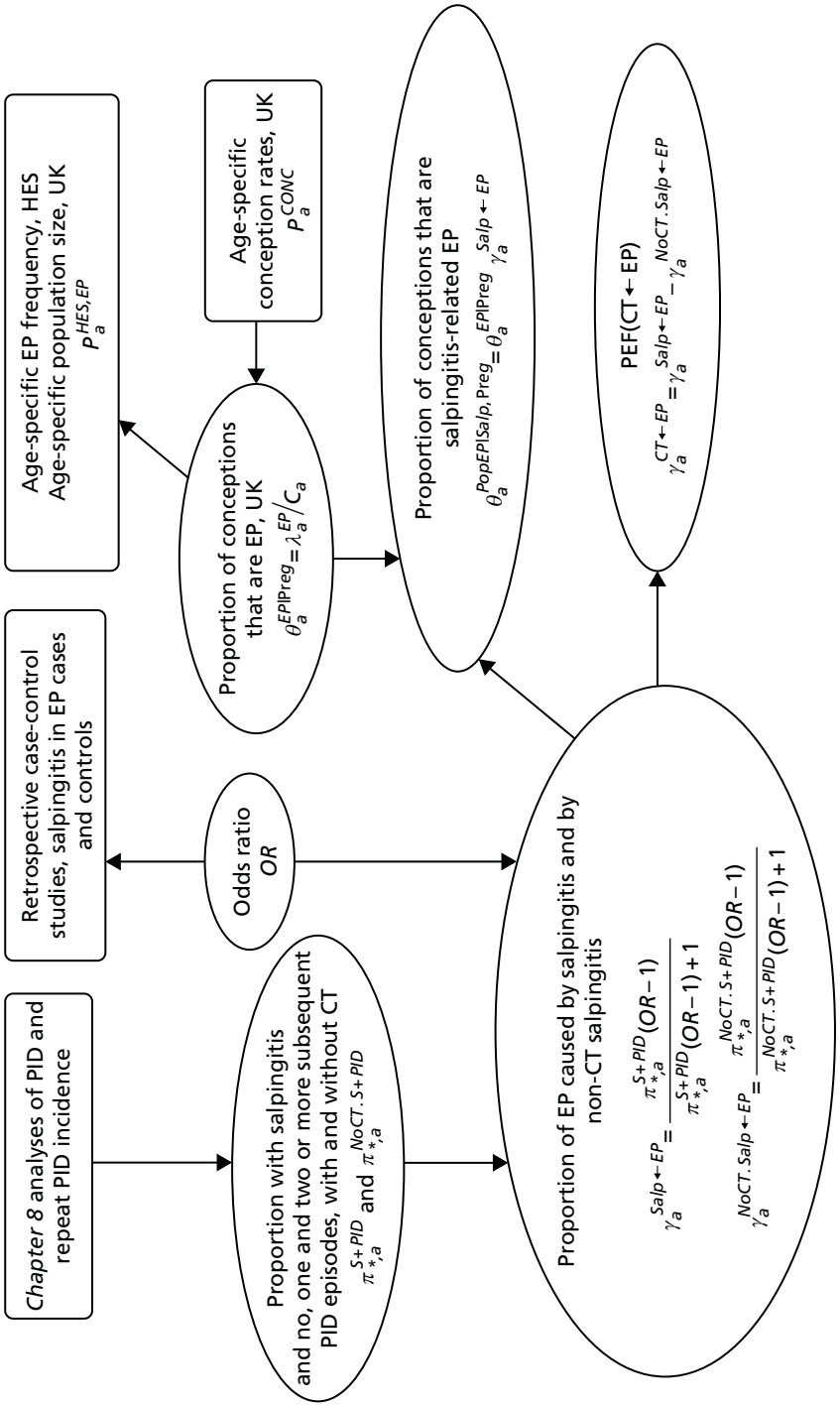

FIGURE 23 Directed acyclic graph of the retrospective analysis.

Results

Risk of ectopic pregnancy in based on prospective data from the Lund study

Table 32 shows the full conditional estimates $\eta_{m,s,a}$ from the model by number of subsequent PID episodes, age group and severity. It also shows the average for hospital diagnosed salpingitis (i.e. averaged over severity with numbers of mild, moderate and severe salpingitis observed in the Lund study.[124] The assumption of no interaction yields a set of cell-specific estimates (where cell is defined as a combination of age, severity, and number of episodes) with an extraordinary range of risk, ranging over 40-fold in the 16- to 24-year-old group from 0.017 for a single 'mild' PID, to 0.698 in a women with three or more subsequent PID episodes following a 'severe' index salpingitis.

The adequacy of these estimates can be judged only from the close agreement of the marginal estimates of the probability that a pregnancy is ectopic by number of subsequent PID episodes, age group and severity, from our model (see *Table 28*, column 3) compared with the crude estimates from the Lund study[124] (see *Table 28*, column 2). The residual deviance was 7.2 for a model with seven parameters and nine data points. The very good fit of the model is to be expected, given the number of parameters and the number of observations.

Table 33 shows the proportions $\pi_{m,d,a}^{S+PID}$ and $\pi_{m,d,a}^{NoCT\ S+PID}$ of the female population at each level of clinical severity of the index episode of salpingitis and the subsequent PID episodes, for the 16–24 and 25–44 age bands, for all salpingitis and non-CT-related salpingitis, respectively.

Population Excess Fraction: the contribution of salpingitis and Chlamydia trachomatis *to ectopic pregnancy*

The upper and lower bounds on the PEF of EP due to salpingitis, PEF (Salpingitis←EP) are given in *Table 34*, by age, based on the PEF formula (see equation 1). Also shown for completeness are the cumulative PID prevalence figures used in the calculation and the ORs. Averaging over the age range 16–44 years, between 21% and 33% of EP is due to salpingitis, these being the lower and upper estimates, with CrIs spanning 11% to 46%. We can take the mean, 27%, as a best estimate. In the absence of CT, the equivalent overall estimates would be 22%, with CrIs spanning 7% to 43%. The proportion of EP attributable to CT is found by subtraction: 4.9% (95% CrI 1.2% to 12.1%).

TABLE 32 Probability a pregnancy is ectopic by age, severity of salpingitis, and number of subsequent PID episodes

Risk factor	Control[b]	No further PID episodes (95% CrI)	One further PID episode (95% CrI)	Two or more further PID episodes (95% CrI)
Age 16–24 years[a]	0.010			
Mild	–	0.017 (0.007 to 0.033)	0.052 (0.019 to 0.103)	0.183 (0.066 to 0.351)
Moderate	–	0.035 (0.021 to 0.053)	0.101 (0.052 to 0.169)	0.316 (0.154 to 0.514)
Severe	–	0.162 (0.117 to 0.211)	0.371 (0.236 to 0.522)	0.698 (0.492 to 0.857)
Overall		0.054 (0.042 to 0.067)	0.139 (0.087 to 0.200)	0.350 (0.209 to 0.505)
Age 25–44 years	0.026			
Mild	–	0.040 (0.015 to 0.076)	0.113 (0.041 to 0.224)	0.335 (0.133 to 0.582)
Moderate	–	0.078 (0.043 to 0.123)	0.207 (0.105 to 0.341)	0.509 (0.280 to 0.733)
Severe	–	0.310 (0.203 to 0.434)	0.572 (0.384 to 0.750)	0.836 (0.664 to 0.943)
Overall		0.112 (0.076 to 0.153)	0.250 (0.153 to 0.361)	0.517 (0.329 to 0.705)

a Assume lower age group applies to 16- to 29-year-olds for the time of the EP.
b Assumes the same age ratio in the control group as the general population. There is no severity distribution in women without salpingitis (controls).
Cell-specific estimates, $\eta_{m,s,a}$ derived from the logistic regression model of the marginal data in *Table 28*.

TABLE 33a Proportion of the entire population in each risk group by age at index salpingitis, diagnosis category, and number of subsequent episodes of PID, expressed as per cent (95% CrI)

	No. of subsequent PIDs			
Age (years)	0	1	2+	Total
16–24				
Undiagnosed	3.26 (1.68 to 5.37)	0.87 (0.34 to 1.71)	0.31 (0.01 to 0.89)	4.44 (2.16 to 6.24)
Non-hospital diagnosed	1.21 (0.70 to 1.77)	0.30 (0.16 to 0.47)	0.09 (0.03 to 0.23)	1.61 (0.96 to 2.31)
Hospital diagnosed	0.58 (0.33 to 0.88)	0.16 (0.08 to 0.25)	0.06 (0.02 to 0.14)	0.80 (0.47 to 1.19)
Total	5.05 (2.89 to 7.61)	1.33 (0.61 to 2.36)	0.47 (0.12 to 1.24)	6.85 (3.78 to 10.7)
25–44				
Undiagnosed	6.34 (3.39 to 5.37)	2.12 (0.89 to 3.99)	0.97 (0.22 to 2.65)	9.42 (4.67 to 16.1)
Non-hospital diagnosed	1.45 (0.70 to 2.38)	0.47 (0.22 to 0.78)	0.20 (0.06 to 0.46)	2.11 (1.05 to 3.40)
Hospital diagnosed	2.04 (1.11 to 3.25)	0.67 (0.36 to 1.08)	0.29 (0.10 to 0.65)	3.00 (1.70 to 4.67)
Total	9.83 (5.77 to 14.4)	3.26 (1.59 to 5.50)	1.46 (0.41 to 3.66)	14.5 (8.16 to 22.3)
16–44				
Undiagnosed	5.02 (2.64 to 8.02)	1.58 (0.65 to 2.98)	0.68 (0.16 to 1.86)	7.27 (3.58 to 12.4)
Non-hospital diagnosed	1.34 (0.74 to 2.07)	0.39 (0.20 to 0.63)	0.16 (0.05 to 0.36)	1.89 (1.06 to 2.86)
Hospital diagnosed	1.41 (0.79 to 2.19)	0.45 (0.24 to 0.71)	0.19 (0.06 to 0.43)	2.05 (1.19 to 3.11)
Total	7.77 (4.52 to 11.5)	2.42 (1.17.4.13)	1.03 (0.28 to 2.59)	11.2 (8.14 to 17.3)

Table gives estimates of $\pi^{S+PID}_{m,d,a}$.

TABLE 33b Proportion of the entire population in each risk group by age at index not-CT-related salpingitis, diagnosis category, and number of subsequent episodes of PID, expressed as per cent (95% CrI)

	No. of subsequent PIDs			
Age (years)	0	1	2+	Total
16–24				
Undiagnosed	2.42 (0.94 to 4.44)	0.52 (0.13 to 1.22)	0.15 (0.02 to 0.51)	3.08 (1.10 to 6.02)
Non-hospital diagnosed	0.86 (0.40 to 1.38)	0.17 (0.06 to 0.33)	0.04 (0.01 to 0.13)	1.08 (0.48 to 1.76)
Hospital diagnosed	0.44 (0.22 to 0.71)	0.10 (0.04 to 0.18)	0.03 (0.00 to 0.08)	0.57 (0.27 to 0.92)
Total	3.73 (1.63 to 6.30)	0.78 (0.23 to 1.70)	0.22 (0.03 to 0.71)	4.73 (1.93 to 8.44)
25–44				
Undiagnosed	5.41 (2.57 to 9.13)	1.51 (0.47 to 3.22)	0.55 (0.08 to 1.74)	7.47 (3.20 to 13.66)
Non-hospital diagnosed	1.22 (0.57 to 2.03)	0.32 (0.12 to 0.60)	0.11 (0.02 to 0.30)	1.65 (0.75 to 2.78)
Hospital diagnosed	1.74 (0.93 to 2.79)	0.47 (0.20 to 0.84)	0.16 (0.03 to 0.43)	2.37 (1.24 to 3.84)
Total	8.36 (4.47 to 12.98)	2.30 (0.85 to 4.47)	0.82 (0.14 to 2.42)	11.48 (5.64 to 19.11)
16–44				
Undiagnosed	4.12 (1.88 to 7.10)	1.08 (0.33 to 2.35)	0.37 (0.05 to 1.21)	5.58 (2.31 to 10.36)
Non-hospital diagnosed	1.06 (0.53 to 1.71)	0.26 (0.10 to 0.47)	0.08 (0.02 to 0.22)	1.40 (0.68 to 2.29)
Hospital diagnosed	1.18 (0.63 to 1.87)	0.31 (0.13 to 0.55)	0.11 (0.02 to 0.28)	1.59 (0.84 to 2.55)
Total	6.37 (3.28 to 10.08)	1.65 (0.58 to 3.27)	0.56 (0.10 to 1.68)	8.57 (4.07 to 14.49)

Table gives estimates of $\pi^{NoCT,S+PID}_{m,d,a}$.

TABLE 34 Retrospective evidence on EP

Age (years)	OR (95% CI)	Cumulative salpingitis (95% CI)	Cumulative salpingitis, no CT	$PEF^{Sal \leftarrow EP}$ (95% CI)	$PEF^{NoCTSalp \leftarrow EP}$ (95% CI)	$PEF^{CT \leftarrow EP}$ (95% CI)
16–24						
Lower bound	3.39 (1.44 to 7.96)	4.15 (2.27 to 6.51)	2.48 (0.66 to 4.83)	9.11 (4.13 to 16.1)	5.65 (1.37 to 12.0)	3.46 (0.92 to 7.95)
Upper bound	5.42 (2.60 to 11.3)	4.15 (2.27 to 6.51)	2.48 (0.66 to 4.83)	15.44 (8.33 to 24.4)	9.80 (2.69 to 18.8)	5.65 (1.61 to 12.4)
25–44						
Lower bound	3.39 (1.44 to 7.96)	13.9 (7.75 to 21.3)	10.84 (5.28 to 18.1)	24.7 (12.8 to 38.8)	20.4 (9.36 to 34.5)	4.28 (1.23 to 9.20)
Upper bound	5.42 (2.60 to 11.3)	13.9 (7.75 to 21.3)	10.84 (5.28 to 18.1)	37.4 (23.6 to 51.4)	31.8 (17.7 to 47.0)	5.63 (1.62 to 11.9)
16–44						
Lower bound	3.39 (1.44 to 7.96)	11.22 (6.27 to 17.3)	8.57 (4.07 to 14.49)	21.1 (10.58 to 34.0)	16.9 (7.41 to 29.6)	4.13 (1.17 to 8.98)
Upper bound	5.42 (2.60 to 11.3)	11.22 (6.27 to 17.3)	8.57 (4.07 to 14.49)	32.7 (20.02 to 46.2)	27.04 (14.2 to 41.4)	5.70 (1.66 to 12.1)

ORs at upper and lower limits, cumulative prevalence and PEFs (as percentages), by age.

Comparison of prospective and retrospective estimates of the proportion of pregnancies that are ectopic as a result of salpingitis

The overall proportion of pregnancies that are ectopic, shown in the last column of *Table 35*, represents an absolute upper bound on the proportion of pregnancies that can be ectopic and due to salpingitis. Taken over the 16- to 44-year age range, 1.13% of conceptions are ectopic. Predictions of the proportion of pregnancies that are ectopic as a result of salpingitis based on the estimate PEF from retrospective data are given in columns 6 and 7: between 0.24% and 0.36%.

The predictions from the five prospective model range from a minimum 0.38 (Model 5: *Av*, *Mi*, *0*) to a maximum 1.54 (Model 1: *Av*, *Av*, *Av*). This means that, given that the PEF is in the range of 21–33%, the Lund study estimates are consistent with the EP rates for England only under Model 5, that is only if we assume there is no EP risk from undiagnosed salpingitis and the risk following salpingitis diagnosed outside the hospital setting is at the rate seen in 'mild' salpingitis in the Lund study (*Figure 24*).

Note that the steep age gradient seen in the observed HES data is closely mirrored by the age gradient in the predictions from the prospective models. This provides a degree of support for the model of cumulative CT and PID incidence developed in *Chapter 8* and our view that EP risk is determined by cumulative exposure to risk factors.

TABLE 35 Percentages of conceptions that are ectopic as a result of PID

Percentage (95% CrI) of conceptions that are ectopic and due to PID								
Prospective estimates, based on Lund study[124]						Retrospective estimates		Percentage of conceptions that are EP (all-cause) England
Age (years)	Model 1: Av, Av, Av	Model 2: Av, Av, Mi	Model 3: Av, Mi, Mi	Model 4: Av, Mi, Mi/2	Model 5: Av, Mi, 0	Lower bound	Upper bound	
16–19	0.10 (0.05 to 0.18)	0.05 (0.02 to 0.10)	0.03 (0.00 to 0.07)	0.02 (0.00 to 0.05)	0.01 (0.00 to 0.03)	0.01 (0.01 to 0.03)	0.03 (0.01 to 0.04)	0.32 (0.29 to 0.36)
20–24	0.43 (0.20 to 0.81)	0.23 (0.09 to 0.46)	0.15 (0.02 to 0.37)	0.10 (0.02 to 0.26)	0.07 (0.02 to 0.13)	0.09 (0.04 to 0.15)	0.15 (0.08 to 0.22)	0.69 (0.65 to 0.73)
25–34	1.39 (0.62 to 2.63)	0.80 (0.31 to 1.64)	0.62 (0.15 to 1.44)	0.45 (0.14 to 1.00)	0.30 (0.13 to 0.55)	0.23 (0.12 to 0.37)	0.36 (0.22 to 0.51)	1.08 (1.04 to 1.11)
35–44	2.57 (1.10 to 4.90)	1.53 (0.54 to 3.20)	1.34 (0.37 to 2.99)	1.01 (0.33 to 2.17)	0.70 (0.30 to 1.26)	0.45 (0.24 to 0.69)	0.66 (0.43 to 0.89)	1.62 (1.54 to 1.69)
16–44	1.54 (0.68 to 2.90)	0.90 (0.34 to 1.85)	0.75 (0.20 to 1.68)	0.56 (0.18 to 1.20)	0.38 (0.17 to 0.68)	0.24 (0.12 to 0.38)	0.36 (0.22 to 0.51)	1.13 (1.10 to 1.16)

The last column is the percentage of conceptions that are ectopic from any cause, based on HES data. It represents an absolute upper limit of the proportion of conceptions that could be due to salpingitis-related EP. Upper and Lower estimates are based on PEFs derived from case–control studies, applied to the England rate. For prospective estimates see text.

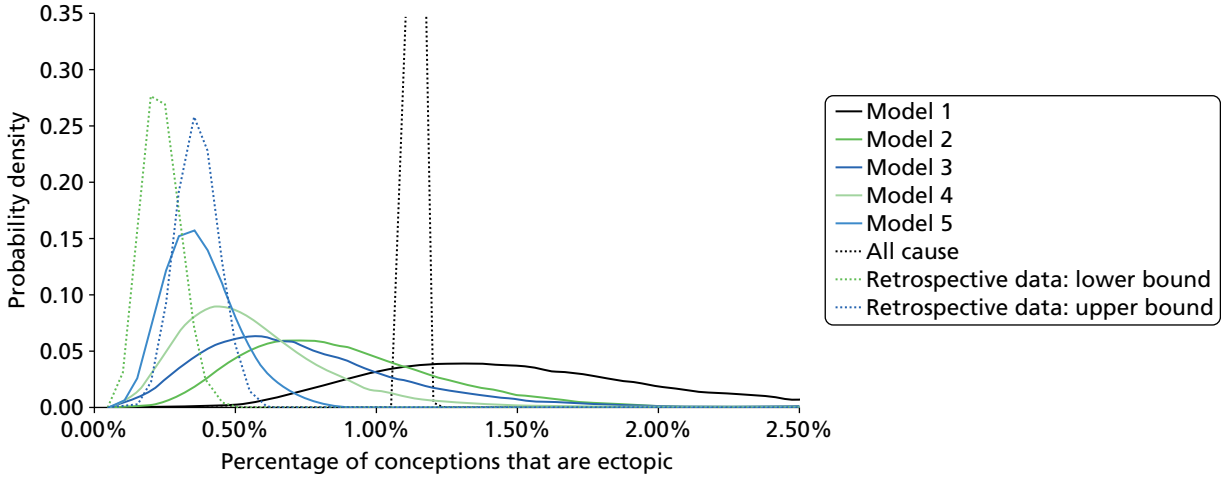

FIGURE 24 Predicted proportions of conceptions that are ectopic due to salpingitis based on five prospective models of the Lund data, and retrospective estimates.

Discussion

Summary of present findings and relation to previous work

The overall proportion of pregnancies that are EP in women aged 16–44 years, based on HES data, is around 1.13%. Based on retrospective studies on the prevalence of a history of salpingitis in EP versus non-EP control subjects, we have suggested that the proportion of EP attributable to PID is between 21% and 33%, so that the proportion of pregnancies that are PID-related EP should be between 0.24% and 0.36%. However, these results are only compatible with progression rates from the Lund study if we assume that there is no EP risk following salpingitis associated with undiagnosed PID, and a level of risk only at the 'mild' level for all salpingitis associated with PID that is diagnosed outside of hospital. Further discussion of these results is deferred to the summary (see *Chapter 12*), where they can be considered alongside results from other chapters, and particular alongside similar analyses of the role of PID/salpingitis in TFI.

Comparison of our results with other studies is problematic because of the different methodologies used, the different assumptions being made about CT and PID, and different ways of taking into account that there must be a pregnancy for there to be an EP. Roberts et al.,[2] for example, assumed that the risk that a future pregnancy would be an EP was 0.008 per day of CT duration, but this was based on a calibration against the Uppsala data, where women with CT were treated, and on CT duration of 200 days, compared with our estimate which is approximately double that.

A slightly different calculation based on similar assumptions was proposed by Low et al.[15] But in both cases the EP rate was assumed to be independent of the PID rate, which runs counter to what is known about their relationship. Adams et al.[106] assumed the risk of EP following PID (conditional on conception) was 7.6%, based on the Lund study,[124] equating to a risk of 0.76% per incident CT infection, given a 10% progression rate from CT to PID. Van Valkengoed *et al.*,[39] based on a form of argument critiqued in *Appendix 9*, proposed that the risk of EP following CT was 10 times less, at 0.07%. Comparison of these results with predictions from this report are reserved till *Chapter 12*, at which point we will have established a credible estimate for the proportion of conceptions that are EP due to PID.

Limitations

There are a wide range of limitations, many of which have already been implicit in the description of our methodology. Undoubtedly, the main difficulty has been the complete lack prospective data on the sequelae of undiagnosed PID, let alone salpingitis following undiagnosed PID, and our reliance on progression rates observed in the Lund study,[124] which was carried out between 30 and 50 years ago, at a time when the microbiological spectrum of PID was more weighted towards gonorrhoea and treatment for PID was quite different. There have been other studies following PID cases forward (see *Table 27*), but these have lacked controls, have been very considerably smaller and have not generally produced different results.

The derivation of predictions from our prospective model also involved a wide range of assumptions that are based on no or very limited evidence. At a technical level our analysis of the separate risk factor-specific risks is not quite correct, as we have taken the age, severity and number of PID episodes as independent data samples. Methods have been developed for deriving estimates from marginal data,[249,250] but these require full conditional data from at least one study. However, *Table 28* shows that our estimates accord well with the evidence available; the precision is overstated as a result of our assumptions about the data, but probably not by very much. Our assumption that there are no interactions between the risk factors cannot be verified with these data.

Also difficult to verify are our assumptions that the severity distribution of any salpingitis remains the same following first, second and third episodes of PID, and that the proportion of clinical PID episodes that are salpingitis is the same in undiagnosed and non-hospital-referred PID as hospital PID.

Our estimates of the PEFs relating to salpingitis and non-CT-related salpingitis must be considered only approximate because the distribution of salpingitis exposure would be different in the absence of CT. However, the use of these retrospective studies is a much greater cause for concern due to the general difficulties with case–control designs (recall bias, confounding), and the uncertain applicability of these results from France to the UK in the present day.

The interpretation of the results in *Table 35* comparing retrospective and prospective results will be discussed further in *Chapters 10* and *12*, where they will be considered alongside similar analyses of TFI.

Summary of assumptions and findings

Summary of assumptions

1. The proportion of all PID, whether diagnosed or not and whether CT related or not, in which salpingitis would be found on laparoscopy, is the same.
2. The severity distribution of salpingitis observed in the Lund study holds for all salpingitis, and whether CT-related or not.
3. The proportion of EP that is due to salpingitis, estimated in a French study in 2001, approximates the proportion in the UK.

Summary of findings

1. Of UK pregnancies, 1.13% are ectopic.
2. An estimated 27% (95% CrI 11% to 46%) of EP in women aged 16–44 years is due to salpingitis.
3. An estimated 4.9% (95% CrI 1.2% to 12.1%) of EP in women aged 16–44 years is due to CT.
4. These results are compatible with the Lund study[124] if it is assumed that there is no EP risk following salpingitis in women with undiagnosed PID, and that in women with salpingitis associated with PID diagnosed outside of hospital, the EP risk is at or below the rate for ‘mild’ salpingitis.

Chapter 10 Pelvic inflammatory disease and tubal factor infertility

Objectives

To:

1. estimate the proportion of women in the UK with TFI at the end of their reproductive lives
2. apply the information on group-specific risks of TFI reported in the Lund study (number of PID episodes, severity of PID) to estimates of the distribution of these risk factors in the UK to predict the proportion of proportion of women who have TFI
3. compare the predictions based on the Lund study with estimates of the proportion of women with TFI derived from UK fertility surveys
4. estimate the proportion of TFI attributable to CT.

Introduction

This chapter examines the relationship between PID, and particularly salpingitis, and TFI. The structure of the chapter and the methods used mirror *Chapter 9* on EP. We can carry over a substantial part of the analysis from the previous chapter, including the estimates of the UK distribution of risk factors (numbers of previous episode of salpingitis, by severity and age: see *Table 33*). As with EP, we use the Lund study[33] on risk of TFI in women with salpingitis who were hospitalised for PID, to define the risk of TFI in the specific subgroups. We then apply the five prediction models that map the Lund study risks in women with hospitalised PID to what might be expected in other diagnosed salpingitis and in undiagnosed salpingitis.

These predictions are then compared with estimates of the prevalence of TFI in the UK, based on fertility surveys. As noted in *Chapter 2*, in order to achieve a definition of TFI prevalence that allows us to compare predictions from the Lund study[33] to prevalence in the UK, we adopt the strategy of considering infertility, including both primary and secondary, as an unresolved failure to conceive at the end of a woman's reproductive life, defined here as age 44 years.

The particular point of interest is whether the prediction model that agrees best with prospective evidence on EP is also the best fit for TFI. This will be examined in *Chapter 12*.

Outline of evidence sources and approach to synthesis

Prospective evidence: following pelvic inflammatory disease forward to tubal factor infertility

A number of prospective studies[228,240–242,251–254] have followed up women with PID to infertility-related outcomes. However, there are major differences in the way that exposures and outcomes have been defined. Studies have recruited women with mild/moderate PID, with severe PID, with all except the most severe PID, or with any PID which is sometimes subclassified to 'mild', 'moderate' or 'severe'. PID was clinical or with laparoscopically confirmed salpingitis, whereas the outcome was either laparoscopically confirmed TFI, unilateral or bilateral tubal occlusion, both, pregnancy rates, or all-cause infertility defined as either 1 or 2 years of trying for a baby, and this could have been primary or all infertility. Finally, a number of studies are based on the same sample. Because of these difficulties, we do not consider these studies further.

Instead, as with EP, rather than attempting to synthesise these disparate studies, we have relied on the Lund study[33] because of its size, length of follow-up, relatively specific exposures and outcomes, and because it uses a similar set of definitions in the follow-up to EP.[124] *Table 36* shows the number of women diagnosed with TFI during follow-up by age group at index salpingitis (< 25 years, ≥ 25 years), its severity and number of subsequent PID episodes during follow-up.

Prevalence of infertility and tubal factor infertility in the UK

The types of evidence sources used to estimate the prevalence of infertility and TFI in the UK were reviewed in *Chapter 2*, and definitions of these terms were given. We identified papers on the prevalence of infertility in the UK and on the prevalence of TFI (*Table 37*), relying on the recent review and synthesis paper by Kavanagh[256] and on references found in our wider review of the literature.

TABLE 36 Prospective data on the risk of TFI following PID from the Lund study[33] by number of PIDs, age and severity

	Aged < 25 years at index			Aged > 25 years at index		
Number of PIDs	*r*	*n*	Crude risk, % (95% CrI)	*r*	*n*	Crude risk, % (95% CrI)
One	59	771	7.7 (6.0 to 9.7)	20	220	9.1 (6.0 to 13.7)
Mild	2	241	0.8 (0.3 to 3.0)	0	71	0.0 (0.0 to 5.1)
Moderate	23	361	6.4 (4.3 to 9.4)	5	89	5.6 (2.5 to 12.6)
Severe	34	369	9.2 (6.7 to 12.6)	15	60	25.0 (16.0 to 37.8)
Two	29	158	18.4 (13.2 to 25.3)	7	27	25.9 (13.7 to 46.2)
Three plus	23	61	37.7 (27.0 to 51.0)	3	4	75.0 (39.7 to 99.4)
Total	111	990	11.2 (9.4 to 13.3)	30	251	12.0 (8.5 to16.7)

r, numerators; *n*, denominators.

TABLE 37 Data used to estimate the prevalence of TFI in women aged 44 years

Publication	Age (years)	*r*	*n*	*p* = *r*/*n*(%) (95% CrI)
Primary infertility				
Templeton[126]	46–50	27	766	3.5 (2.4 to 5.1)
Gunnell[128]	41–50	31	1609	1.9 (1.4 to 2.7)
Bhattacharya[127]	41–50	79	2347	3.4 (2.7 to 4.2)
Oakley[129]	40–55	159	6584	2.4 (2.1 to 2.8)
Oakley, adjusted[129]	40–55	153.8	6128	2.5 (2.2 to 2.9)
Not involuntarily childless				
Templeton[126] and Bhattacharya[127]	41–50	2910	3113	93.5 (92.6 to 94.3)
Secondary infertility				
Templeton[126]	46–50	17	766	2.2 (1.4 to 3.5)
Gunnell[128]	41–50	41	1609	2.5 (1.9 to 3.4)
Bhattacharya[127]	41–50	5	2347	0.2 (0.1 to 0.5)
Proportion TFI				
Maheshwari[255]	> 35	442	1782	24.8 (22.9 to 26.9)

n, denominator; *p* = study-specific estimate; *r*, numerator.

The approach we adopt takes the following steps:

1. The data on primary infertility in women aged 41–50 years from four surveys[126–129] is pooled. The data from one of the surveys[129] has to be adjusted to take account of the proportion of women who were voluntarily childless, which was estimated from data in two surveys.[126,127]
2. A second synthesis was carried out for the prevalence of secondary infertility, using data from three surveys,[126–128] and this estimate was adjusted for the possibility that these women would have become pregnant subsequently (see *Appendix 13*).
3. The primary infertility and adjusted secondary infertility estimates were added together to form an estimate of the prevalence of infertility.
4. This was multiplied by an estimate of the proportion of infertility due to TFI. Because data from the HFEA are unlikely to be representative, we have used an estimate of the proportion of infertility due to TFI of approximately 25%, based on women aged > 35 years old from a clinic-based study conducted in Aberdeen between 1993 and 2006.[255] This is the only UK survey of infertility reporting this statistic. Approximately 13% of tubal damage can be caused by endometriosis[257] but this is classified as endometriosis rather than TFI in that study.

As noted in *Chapter 2*, we assume that virtually all TFI has an infective aetiology and is caused by salpingitis, and that PID that does not confirm as salpingitis on laparoscopy does not carry a risk of TFI.[35]

Statistical methods

Model for the prospective data of the risk of tubal factor infertility following pelvic inflammatory disease

The Lund study[33] provides data on the risk of TFI first by age and severity of index salpingitis, and second by age and number of subsequent PID episodes, zero, one and two-plus, averaging over severity (see *Table 36*). Each of the 10 data points describing progression to TFI has a binomial likelihood,

$$\begin{aligned} r_{1,s,a} &\sim Binomial(\eta_{1,s,a}, n_{1,s,a}) \\ r'_{m,a} &\sim Binomial(p'_{m,a}, n'_{m,a}) : m = 2,3 \end{aligned} \tag{58}$$

where a indexes over age group (< 25 or > 25 years at recruitment), m over number of subsequent PID episodes during follow-up (zero, one, two-plus) and s over severity of the index salpingitis episode (s = 1, mild; 2, moderate; 3, severe). $r_{1,s,a}$ is the number of women with one salpingitis episode of severity s at age a developing TFI in a sample size $n_{1,s,a}$ of women in these groups who were follow-up and tried to conceive. $r'_{m,a}$ is the number of women with one or two subsequent PID episodes by age group (aggregated over severity) in a sample size of $n'_{m,a}$ with probability parameter $p'_{m,a}$ where:

$$p'_{m,a} = \frac{\sum_{s=1}^{3} \eta_{m,s,a} \cdot n_{1,s,a}}{\sum_{s=1}^{3} n_{1,s,a}} \quad : m = 2,3 \tag{59}$$

As with EP, we fit a model to the marginal data to estimate the risk factor-specific risk of TFI, $\eta_{m,s,a}$:

$$logit(\eta_{m,s,a}) = \beta_0 + \beta_{1,m}M + \beta_{2,s}S + \beta_{3,a}A \quad (60)$$

β_0 is the log odds of TFI in a woman aged < 30 years at the time of her mild index salpingitis, with no further PID episodes, $\beta_{1,m}$ is the LOR of TFI for a woman with *m* subsequent PID episodes compared with none, $\beta_{2,s}$ is the LOR for a woman whose index salpingitis was moderate or severe compared with mild and $\beta_{3,a}$ is the LOR in women aged 30 years or over relative to those aged under 30 years. The covariates *M*, *S* and *A* have the same interpretations as in *Chapter 9*. The parameters β_0, $\beta_{1,m}$, $\beta_{2,s}$ and $\beta_{3,a}$ are given flat normal priors. As with our model in *Chapter 9*, the reported data do not provide the sufficient statistics to identify interaction terms. Although this does not necessarily mean there are no interactions, we assume there are none, as in the analysis of EP. We assume that results for 16- to 24-year-old women are applicable for women aged 16–29 years because of the mean 9-year follow-up.

In the Lund study,[33] TFI was not found in any of 601 control women with PID which was not confirmed as salpingitis on laparoscopy.

Model for the distribution and diagnostic setting of pelvic inflammatory disease in the general population.

Estimates of the proportion of women in age band *a* who had experienced an episode of salpingitis at age *a*, diagnosed in setting *d*, and followed by *m* subsequent PID episodes, $\pi_{m,d,a}^{S+PID}$ and $\pi_{m,d,a}^{NoCT,S+PID}$ were introduced in *Chapter 9* (see *Table 33*).

Prospective estimate of the prevalence of tubal factor infertility caused by CT-related PID

To estimate the prevalence of TFI due to salpingitis in the UK $\theta_a^{POP,TFI|salp}$, we use the same approach as in *Chapter 9* for EP. We adopt the same five 'models' to 'map' the $\eta_{m,s,a}$ severity-specific estimates of TFI risk into $\theta_{m,d,a}^{TFI|salp}$ diagnostic-setting specific estimates of risk. Then, the population estimates $\theta_a^{POP,TFI|salp}$ are found by taking a weighted average of the $\theta_{m,d,a}^{TFI|salp}$ using the $\pi_{m,d,a}^{S+PID}$ as weights.

As the Lund study[33] was set in the Hospital Diagnosed setting, we can estimate $\theta_{m,HD,a}^{TFI|salp}$ directly from the $\eta_{m,s,a}$. Note that, because there were no TFI cases among the non-salpingitis controls, we interpret the parameters $\theta_{m,d,a}^{TFI|salp}$ as causal estimates:

$$\theta_{m,HD,a}^{TFI|salp} = \frac{\sum_{s=1}^{3} \eta_{m,s,a} \cdot n_s^{Lund}}{\sum_{s=1}^{3} n_s^{Lund}} \quad (61)$$

Then, the risk of TFI due to salpingitis as a function of age group, severity and number of subsequent PID episodes is constructed as a weighted average using the proportion of women at each severity level and the cumulative prevalence of salpingitis and zero, one or two-plus subsequent PID episodes, $\pi_{m,d,a}^{S+PID}$. An analogous estimate can be formulated for salpingitis that is not CT related:

$$\theta_a^{Pop,TFI|salp} = \sum_{d=HD,nHD,UD} \sum_{m=0,1,2+} \pi_{m,d,a}^{S+PID} \cdot \theta_{m,d,a}^{TFI|salp} \quad (62)$$

$$\theta_a^{PopnoCT,TFI|salp} = \sum_{d=HD,nHD,UD} \sum_{m=0,1,2+} \pi_{m,d,a}^{noCT,S+PID} \cdot \theta_{m,d,a}^{TFI|salp} \quad (63)$$

The five different models of how the TFI risks in the Lund hospital-diagnosed PID might apply to non-hospital-diagnosed PID, and undiagnosed PID can then be set out in the same way as in *Chapter 9*, where, as before, '*Av*' refers to the average risk reported in the Lund study[33] and '*Mi*', to the risk in mild salpingitis, with *HD*, *nHD* and *UD* referring to hospital diagnosed, non-hospital diagnosed and undiagnosed PID:

Model 1 (Av, Av, Av): Lund average risk of TFI due to salpingitis following all PID. All salpingitis, whether associated with HD, nHD, or UD PID, carries the same risk of TFI as those with hospital diagnosed PID in the Lund study:

$$\theta^{TFI|salp}_{m,nHD,a} = \theta^{TFI|salp}_{m,UD,a} = \frac{\sum_{s=1}^{3} \eta_{m,s,a} \cdot n_s^{Lund}}{\sum_{s=1}^{3} n_s^{Lund}} \tag{64}$$

Model 2 (Av, Av, Mi): Lund average for salpingitis following diagnosed PID, Lund 'mild' following undiagnosed PID. Salpingitis associated with both *HD* and *nHD* PID carries the same TFI risk as in the Lund study; salpingitis associated with *UD* PID has the same TFI risk as the 'mild' salpingitis group in the Lund study.

$$\begin{aligned} \theta^{TFI|salp}_{m,nHD,a} &= \frac{\sum_{s=1}^{3} \eta_{m,s,a} \cdot n_s^{Lund}}{\sum_{s=1}^{3} n_s^{Lund}} \\ \theta^{TFI|salp}_{m,UD,a} &= \eta_{m,1,a} \end{aligned} \tag{65}$$

Model 3 (Av, Mi, Mi): Lund average for salpingitis following hospital diagnosed, Lund 'mild' following non-hospital diagnosed and undiagnosed salpingitis. In this model HD salpingitis progresses to TFI at the average rate observed in the Lund study; nHD and UD at the Lund 'mild' rate.

$$\theta^{TFI|salp}_{m,nHD,a} = \theta^{TFI|salp}_{m,UD,a} = \eta_{m,1,a} \tag{66}$$

Model 4 (Av, Mi, Mi/2): Lund average for salpingitis following HD PID, Lund 'mild' following nHD PID. Salpingitis following UD PID progresses at a rate that is half-way between the 'mild' rate and zero.

$$\begin{aligned} \theta^{TFI|salp}_{m,nHD,a} &= \eta_{m,1,a} \\ \theta^{TFI|salp}_{m,UD,a} &\sim Uniform(0, \eta_{m,1,a}) \end{aligned} \tag{67}$$

Model 5 (Av, Mi, 0): Lund average for salpingitis following HD PID, Lund 'mild' for following nHD PID, no risk following UD PID:

$$\begin{aligned} \theta^{TFI|salp}_{m,nHD,a} &= \eta_{m,1,a} \\ \theta^{TFI|salp}_{m,UD,a} &= 0 \end{aligned} \tag{68}$$

The WinBUGS code for the prospective analyses of TFI is set out in *Appendix 15*.

Model for the prevalence of tubal factor infertility based on UK fertility surveys

All the data, which have binomial likelihoods, are set out in *Table 37*. The data on primary infertility in women aged 41–50 years is pooled on the assumption that the proportions are from a common beta distribution, whose parameters are estimated from the binomial data, and are given vague exponential prior distributions. The target parameter is the mean of this distribution. The data from one of the surveys[129] is adjusted to take account of the proportion of women who were voluntarily childless. This fraction was estimated from data in two surveys.[126,127] The effective sample size of the Oakley survey was imputed based on a simulation taking account of the uncertainty in both fractions.

The synthesis of secondary infertility, based on three surveys, is also based on a common beta distribution, and adjusted for over-reporting, which was assumed to be between 0% and 12% (see *Appendix 13*). Total infertility was multiplied by an estimate of the proportion of infertility due to TFI,[255] based on a vague Beta prior and binomial likelihood. WinBUGs code for the synthesis is provided in *Appendix 14*.

The proportion of tubal factor infertility episodes caused by chlamydia trachomatis

Under our assumption that all TFI is caused by salpingitis, we can estimate the proportion of TFI due to CT, the PEF, by age group, as a function of the prevalence of TFI due to salpingitis, and the predicted prevalence of TFI due to salpingitis if there is no CT. These are based on the estimates of $\pi_{m,d,a}^{S+PID}$ and $\pi_{m,d,a}^{nonCT\ S+PID}$ from *Chapter 8*, through equations 61 and 62, respectively:

$$\gamma_a^{CT \leftarrow TFI} = \frac{\theta_a^{POP,TFI|salp} - \theta_a^{POPnoCT,TFI|salp}}{\theta_a^{POP,TFI|salp}} \tag{69}$$

The same estimate will be derived from each of the five prospective models. A DAG of the prospective analysis can be found in *Figure 25*.

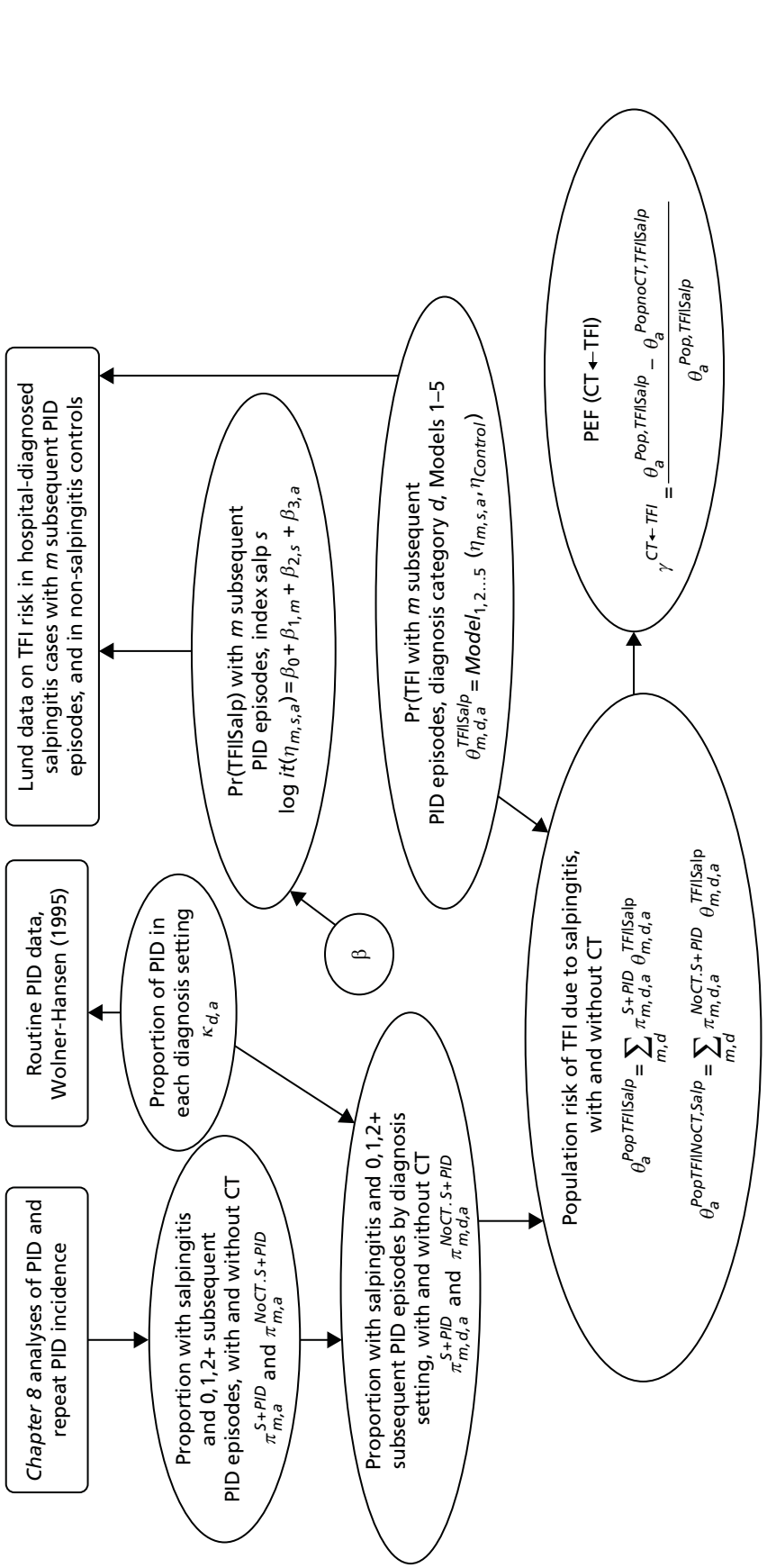

FIGURE 25 Directed acyclic graph of the prospective analysis of TFI.

Results

Population prevalence of tubal factor infertility based on infertility surveys

The posterior summaries from our synthesis of the British fertility surveys are shown in *Table 38*. The prevalence of infertility in women at the end of their reproductive lives, notionally age 44 years for the purposes of this project, including both primary and secondary infertility, is 4.4% (95% CrI 3.2% to 6.2%) and the prevalence of TFI is 1.1% (95% CrI 0.8% to 1.6%).

Risk of tubal factor infertility following salpingitis, by age, severity and number of exposures

Our modelled cell-specific risks $\eta_{m,s,a}$ of TFI, based on the marginal risk table in the Lund study,[33] are set out in *Table 39*. The risk of TFI in women aged 16–29 years is only slightly less than in women aged 30–44 years, but there is a 20-fold increase in risk as one goes from mild to severe disease, and a more than threefold increase from no subsequent PID episodes, to one and then two or more episodes, suggesting a positive multiplicative relationship between number of salpingitis episodes and risk of TFI.

TABLE 38 Estimates of primary, secondary infertility and proportion TFI: posterior means and 95% CIs

Parameter	Mean (%)	95% CI
Primary infertility	2.82	2.25 to 3.50
Secondary infertility	1.63	0.66 to 3.44
Secondary infertility, adjusted	1.54	0.63 to 3.26
All infertility	4.36	3.23 to 6.19
Prevalence of TFI	1.08	0.79 to 1.55

TABLE 39 Risk of TFI due to PID, by age, severity of index salpingitis and number (95% CrI) of subsequent PID episodes

		PID episode(s)		
Risk factor	No salpingitis	No further	One further	Two further
16–24				
Mild	–	0.01 (0.00 to 0.02)	0.02 (0.00 to 0.06)	0.07 (0.01 to 0.18)
Moderate	–	0.06 (0.04 to 0.08)	0.17 (0.11 to 0.25)	0.44 (0.28 to 0.63)
Severe	–	0.21 (0.15 to 0.27)	0.46 (0.34 to 0.60)	0.76 (0.61 to 0.88)
Overall	0			
25–44				
Mild	–	0.01 (0.00 to 0.02)	0.02 (0.00 to 0.07)	0.08 (0.01 to 0.23)
Moderate	–	0.07 (0.04 to 0.11)	0.21 (0.11 to 0.33)	0.49 (0.29 to 0.71)
Severe	–	0.24 (0.16 to 0.33)	0.51 (0.35 to 0.67)	0.79 (0.63 to 0.91)
Overall	0			

Cell-specific estimates of $\eta_{m,s,a}$ derived from the logistic regression model of the marginal data in *Table 36* (mean 8.9-year follow-up).
Assumes lower age group applies to 16- to 29-year-olds for the time of the TFI.
There is no severity distribution in controls.

Comparison of tubal factor infertility prevalence predicted from the Lund study with tubal factor infertility in British surveys

The 1.08% (95% CrI 0.79% to 1.54%) prevalence of TFI in 44-year-old women, based on the above synthesis of the British fertility surveys can now be compared with the range of projected TFI prevalence obtained when the Lund study progression rates from salpingitis to TFI[33] are applied to the distribution of age, severity of index salpingitis, and number of subsequent PID episodes (*Table 40* and *Figure 26*). The estimates for age 44 years appear to fall between the Model 1 (*Av*, *Av*, *Av*) projection of 1.74% (95% CrI 0.9,2.9) (all salpingitis progresses at the average rate observed in the Lund study), and the Model 2 (*Av*, *Av*, *Mi*) projection of 0.70% (95% CrI 0.4% to 1.2%) (all diagnosed salpingitis progresses at the average rate, undiagnosed at the 'mild' rate). The estimates are not compatible with the Model 3 (*Av*, *Mi*, *Mi*) projection of 0.5% (95% CrI 0.2% to 0.9%), even taking into account the uncertainty in the estimates and the projections.

TABLE 40 Prevalence (95% CrI) of TFI as predicted from the distribution of salpingitis severity and exposure, and Lund study progression rates[33] (prospective models 1–5), compared with prevalence of TFI in England, based on fertility surveys

	Prospective model					
Age (years)	1 (Av, Av, Av)	2 (Av, Av, Mi)	3 (Av, Mi, Mi)	4 (Av, Mi, Mi/2)	5 (Av, Mi, 0)	All TFI in England
16–19	0.14 (0.07 to 0.23)	0.06 (0.03 to 0.09)	0.02 (0.01 to 0.04)	0.01 (0.01 to 0.03)	0.01 (0.01 to 0.02)	–
20–24	0.44 (0.23 to 0.73)	0.18 (0.10 to 0.28)	0.06 (0.03 to 0.14)	0.05 (0.02 to 0.11)	0.04 (0.02 to 0.07)	–
25–34	1.00 (0.52 to 1.64)	0.40 (0.22 to 0.63)	0.21 (0.10 to 0.38)	0.19 (0.09 to 0.32)	0.16 (0.09 to 0.26)	–
35–44	1.57 (0.78 to 2.67)	0.63 (0.33 to 1.04)	0.45 (0.21 to 0.81)	0.41 (0.20 to 0.70)	0.37 (0.18 to 0.65)	–
44	1.74 (0.88 to 2.94)	0.70 (0.36 to 1.16)	0.50 (0.23 to 0.89)	0.46 (0.22 to 0.80)	0.41 (0.17 to 0.53)	1.08 (0.79 to 1.54)
16–44	1.03 (0.53 to 1.71)	0.41 (0.22 to 0.66)	0.26 (0.13 to 0.45)	0.23 (0.12 to 0.39)	0.21 (0.11 to 0.34)	–

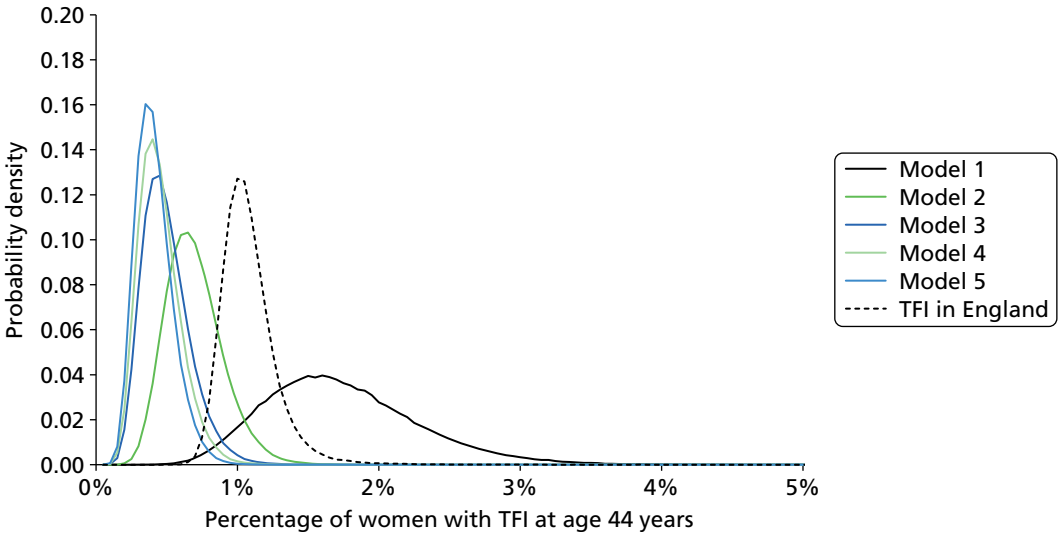

FIGURE 26 Predicted prevalence of TFI in 44-year-olds, based on five prospective models of the Lund data.[33]

*Estimates of the population excess fraction (*Chlamydia trachomatis←*tubal factor infertility) based on the prospective analysis*

Table 41 compares the predicted population prevalence of TFI under Models 1 and 2, as presented in *Table 40*, with alternative predictions under a scenario in which CT-related salpingitis did not exist. A prediction for TFI prevalence caused by CT-related salpingitis is formed by subtraction, leading to an estimate of the PEF(CT←TFI) of 28.7% (95% CrI 8.9% to 56.3%). Note that this estimate of the PEF does not depend on the model for TFI risk that is adopted.

TABLE 41 Estimates of the percentage (95% CrI) of the female population with TFI caused by all salpingitis, with TFI caused by non-CT-related salpingitis, and CT-related salpingitis, and estimates of the PEF(CT←TFI)

Age (years)	Prevalence of TFI due to: Salpingitis A	Non-CT-related salpingitis B	CT-related salpingitis A–B	PEF(CT←TFI), (A–B)/A (as %)
Model 1 (Av, Av, Av)				
16–19	0.14 (0.07 to 0.23)	0.07 (0.00 to 0.16)	0.07 (0.02 to 0.16)	52.7 (16.1 to 100)
20–24	0.44 (0.23 to 0.73)	0.28 (0.08 to 0.54)	0.17 (0.05 to 0.36)	38.1 (11.9 to 74.5)
25–34	1.00 (0.52 to 1.64)	0.75 (0.33 to 1.33)	0.25 (0.08 to 0.54)	25.8 (8.1 to 50.5)
35–44	1.57 (0.78 to 2.67)	1.26 (0.57 to 2.28)	0.31 (0.08 to 0.67)	20.0 (5.9 to 40.0)
44	1.74 (0.88 to 2.94)	1.41 (0.61 to 2.54)	0.33 (0.10 to 0.71)	19.2 (6.0 to 38.0)
16–44	1.03 (0.53 to 1.71)	0.79 (0.35 to 1.42)	0.24 (0.07 to 0.51)	28.7 (8.9 to 56.3)
Model 2 (Av, Av, Mi)				
16–19	0.06 (0.03 to 0.09)	0.03 (0.00 to 0.06)	0.03 (0.01 to 0.07)	52.7 (16.1 to 100)
20–24	0.18 (0.10 to 0.28)	0.11 (0.04 to 0.20)	0.07 (0.02 to 0.15)	38.1 (11.9 to 74.5)
25–34	0.40 (0.22 to 0.63)	0.30 (0.14 to 0.50)	0.10 (0.03 to 0.22)	25.8 (8.1 to 50.5)
35–44	0.63 (0.33 to 1.04)	0.50 (0.24 to 0.87)	0.12 (0.03 to 0.28)	20.0 (5.9 to 40.0)
44	0.70 (0.36 to 1.16)	0.56 (0.27 to 0.97)	0.13 (0.04 to 0.30)	19.2 (6.0 to 38.0)
16–44	0.41 (0.22 to 0.66)	0.31 (0.15 to 0.54)	0.10 (0.03 to 0.21)	28.7 (8.9 to 56.3)

Discussion

In this section we comment on the relation between our results in this chapter on risks of TFI following salpingitis, in the light of previous work on this topic. We then turn to consider the relationship between the results in this chapter and those from *Chapter 9*, where the observed UK EP rates were compared with the same five models of prospective risk derived from the Lund study.

Comparison with previous work on Chlamydia trachomatis*, pelvic inflammatory disease and tubal factor infertility*

We have already provided some comments on the relation between CT, PID and TFI in the HPA and ClaSS economic models, and on the approach taken by van Valkengoed *et al.*[39] (see *Appendix 9*). There appear to have been few attempts to reconcile information on incidence and prevalence of CT, PID and TFI in a quantitative way since then, until the important recent work by Kavanagh *et al.*[256] This attempts to estimate the risk of TFI associated with CT infection. This study does not use PID or salpingitis as an intermediate variable, but, instead, considers only previous CT. For women aged 40–44 years (estimates are also provided for younger age groups) they estimated a prevalence of TFI of 0.75%, that 43% of women in this age group in the general population have ever had chlamydia, and that, based on serological surveys of women with TFI, 72% have ever had chlamydia. From this the authors estimate the probability of TFI in women ever infected with chlamydia to be 1.3%, and estimate a PEF for CT of 62.9%.

Although there are similarities with our estimates, there are several difficulties with Kavanagh *el al*.'s approach, many of which are highlighted in their paper: first, the estimated prevalence of TFI at age 40–44 years (0.75%) may be too low, as only primary infertility was considered. Our estimate was 1.1%, which includes both primary and secondary infertility. Second, the estimate of cumulative incidence of CT, which is taken to be 42.9% by age 44 years, based on testing and prevalence data, must be an overestimate because it sums over incidence rates rather than summing over conditional incidence proportions applied to women who have never been infected.

Their estimates of the prevalence of TFI in women with a history of CT are derived by substituting the above parameters into the Bayes' theorem. However, as the authors point out (also see *Appendix 9*), this is not a causal estimate. Even if CT never caused TFI, and CT exposure was not confounded with other STIs, we would still expect this estimate to equal the background prevalence of TFI. As a result of the above issues, Kavanagh *et al.*'s estimates of the proportion of TFI attributable to CT must be considered to be an upper bound, and, on this basis, are not incompatible with our estimated PEF(CT←TFI) of 28.7%.

This estimate, it should be emphasised, although derived from the prospective models based on the Lund study, is entirely independent of the model selected, and is chiefly influenced by the estimates of the CT incidence (see *Chapter 5*), PID incidence and the PEF(CT←PID) (see *Chapter 7*), and the Markov model for repeat infection (see *Chapter 8*).

Changes in pelvic inflammatory disease management since the Lund study

There are a number of reasons for questioning whether the rates of progression of laparoscopically defined PID to EP reported in the Lund study would apply to the UK now.

Antibiotic treatments for PID have certainly changed since the Lund study, with an increasing use of broad-spectrum antibiotics capable of treating the range of micro-organisms that are now implicated in PID.[258] Although these treatments are highly effective in achieving microbiological cure and elimination of PID symptoms, little is known about their effect on reproductive outcomes.

One treatment variable that is known to have a major impact is time to treatment. In the Lund study (1960–84), in which it was reported that 18% were treated within 72 hours, treatment in less than 72 hours from the onset of symptoms was associated with a 2.8-fold reduction in risk of infertility (95% CI 1.3 to 6.1).[107] During the PEACH study (1996–1999),[259] there were still only 23% treated within 3 days. It is generally believed that in the UK, where access to health care is universal, there was a distinct shift in the mid to late 1990s towards presumptive treatment and treatment of milder cases.[260]

If we assume, optimistically, that 50% of diagnosed PID is treated within 3 days, and that the risk reduction is by 2.8, then we would predict that progression rates among diagnosed PID cases, which would confirm on laparoscopy, should be approximately 76.7% of the rates reported in the Lund study. Applying this reduced risk to the prevalence of TFI predicted by the five models in *Table 40*, and bearing in mind that the reduction in progression rates would apply to only the minority *diagnosed* salpingitis, it is evident that changes in treatment efficacy and management do not materially affect our conclusions, other than to shift the balance more in favour of Model 1 (*Av*, *Av*, *Av*).

Relation between predictions for ectopic pregnancy and tubal factor infertility

The main finding from this chapter is that the model of the Lund study that could predict the UK EP rates adjusted down to reflect the proportion of EP caused by salpingitis, namely Model 5 (*Av*, *Mi*, *0*), cannot predict the UK TFI rates, which are compatible with Models 1 (*Av*, *Av*, *Av*) or 2 (*Av*, *Av*, *Mi*) and vice versa. Specifically, the models that were compatible with the EP data would under-predict the TFI results by a factor of 2–2.5, whereas the models compatible with the TFI data would over-predict the EP results by a factor of 1.5–3, if we accept the retrospective estimates of the proportion of EP due to salpingitis. This issue is taken up in *Chapter 12*.

Assumptions and limitations

The evidence on infertility is based on studies undertaken over a 20-year period. However, following our comments on the aetiology of TFI, we expect that TFI prevalence will respond relatively slowly to changes in CT and PID incidence. We therefore assume that our analyses of this data are approximately relevant to the target period for this report, 2002.

We also need to acknowledge assumptions inherent in our interpretation of the Lund study results on progression from PID to TFI.[33] The limited follow-up of the Lund cohort means that, although the study has a relatively long follow-up time (mean 8 years), there is still the possibility that some cases of TFI will not have been identified at the end of the observation period, leading to a potential *under*estimation of the risk of TFI. However, because the denominator is women who have ever been pregnant this should not cause bias. Also, the same women may have gone on to experience further PID episodes leading to TFI at a later date.

A second issue to bear in mind is the small amount of data available on the risk of TFI in those with a single 'mild' salpingitis. Only 2 out of 241 women aged < 25 years at the index PID with a single mild salpingitis developed TFI, and 0/71 women aged ≥ 25 years developed TFI (see *Table 36*). Although this is fairly strong evidence of low risk, this is the largest risk group, and projections for a number of the five models are therefore very sensitive to this particular parameter. If there had been just one observed case of TFI, rather than none, in those aged > 25 years, this would have had a substantial impact on our estimates of $\theta_a^{popTFI|salp}$ in that group.

Summary of assumptions and findings

Summary of assumptions

1. The proportion of PID in which salpingitis would be found on laparoscopy is the same, whether diagnosed or not, and whether CT related or not.
2. The distribution of severity of salpingitis observed in the Lund study holds for all salpingitis, whether diagnosed or not, and whether CT related or not.

Summary of findings

1. The proportion of women in the UK who are infertile at the end of their reproductive lives (age 44 years) is estimated to be 4.4% (95% CrI 3.2% to 6.2%).
2. The proportion with TFI is 1.1% (95% CrI 0.8% to 1.6%).
3. These results are compatible with the Lund study if it is assumed that the TFI risk following salpingitis in women with diagnosed PID is at the average level observed in the Lund study,[33] and among those with undiagnosed PID the rate is between the 'mild' and average rates.
4. The estimated proportion of TFI attributable to CT is 29% (95% CrI 9% to 53%).
5. The prospective model of the Lund data that provides the best fit between observed and predicted TFI prevalence, does not agree with the best-fitting model for EP.

Chapter 11 Tubal factor infertility: previous exposure to *Chlamydia trachomatis* infection

Objectives

To estimate the proportion of TFI attributable to CT using retrospective serological case–control studies.

Introduction

There have been a large number of retrospective studies comparing the prevalence of serum antibodies to CT in women with TFI and controls.[143,261,262] The motivation was to assess the diagnostic potential of serological tests, or to study the causal role of CT in TFI. In some cases the purpose was to estimate the proportion of TFI that could be causally attributable to CT,[263] using the formula for PEF (equation 1) presented in *Chapter 3*.

Besides the issue of positive confounding with other STIs (see *Chapter 3*), the use of this formula in serological studies comparing TFI with controls faces two further obstacles. First, as indicators of previous CT infection, CT antibody tests are relatively insensitive. Their specificity is also poor, with earlier assays showing cross-reactivity with antigens to *Chlamydia pneumoniae*.[262,264] Even so, given estimates of sensitivity and specificity,[55] estimates of true prevalence can be recovered from observed CT prevalence in cases and controls. Unfortunately, reliable estimates of sensitivity and specificity are not generally available. Sensitivity can be assessed in individuals identified as infected following bacterial culture or NAATs. But a true-negative population in which specificity can be estimated is less easy to identify: patients who are culture negative and/or NAAT negative cannot be regarded as 'never infected', as they may have experienced a CT infection which subsequently cleared.[169] 'Discrepancy analysis', as proposed by Morre *et al.*,[265] is recognised to overestimate test accuracy.[266,267] A better strategy to assess specificity is to take samples from children, who would not normally be expected to have been exposed to CT. It is necessary to choose children who old enough to have cleared maternal antibody.[268,269]

A further complication is that antibody levels, and hence test sensitivity, may not be the same in women with TFI whose TFI was caused by CT and women who were previously infected but do not have TFI. There are consistent reports of higher antibody titres in women with TFI[270,271] compared with controls, including control subjects who have been previously exposed to CT. It is believed that these high titres reflect an inflammatory response to CT, which may, in turn, cause tubal occlusion.[272] The same findings have been reported in women with salpingitis and women with EP.[88,273]

The standard formula that allows estimation of the PEF from case–control studies is, therefore, not to be relied on. For example, one study[263] using a micro-immunofluorescence (MIF) assay reported antibody prevalence in TFI cases was 72%, compared with 22%, an OR of 9.2, and an aetiological fraction of 64%. This would imply that 64% of all TFI could be attributed to CT. The high reported prevalence in TFI cases is sometimes attributed to false positives, but sensitivity must always be higher than observed prevalence. For example, if we make an optimistic assumption that sensitivity is 85% and specificity 95%, the corrected prevalences would be 90% and 21.5%, an OR of 33.4 and an aetiological fraction of 87.3%. However, lower levels of sensitivity yield ORs and aetiological fractions that are even more extreme. The study can be explained if it is assumed that CT antibody titre, and hence sensitivity, is higher in the subgroup whose TFI was caused by CT. This hypothesis, suggested above, forms the basis for the approach we have taken.

In this chapter we propose a new way to estimate the population aetiological fraction of TFI due to CT. The method is applied to a comparative study of antibody prevalence in TFI cases and controls, which used three peptide-based assays and two MIF assays.[274] Our re-analysis of this data set addresses the problems attaching to equation 1: we adjust for the bias caused by imperfect sensitivity and specificity of the assays; we take account of the especially high antibody levels in women with TFI caused by CT; and for the confounding influence of other STIs.

We begin by setting out our approach to literature identification, and explaining the reason why have focused attention on the Land *et al.* study.[274] We then show how test sensitivity and specificity can be characterised, and how properties of the tests may depend on the stage that the patient has reached in the natural history of CT infection. We then outline how the Land *et al.* data[274] on the prevalence of anti-CT antibody in women with TFI and control subjects, in several tests, can be re-analysed to estimate of the proportion of TFI that might be due to CT. Some additional analyses and technical details are given in *Appendix 16*. The work in this chapter has been published previously.[275]

Methods

Evidence identification

We identified studies from review papers,[143,261,276] from co-authors' knowledge of the literature, and as part of a wider review of the literature on the natural history of CT.

Although there have been many studies of CT antibody prevalence in TFI cases and controls, the statistical modelling methods used in this paper require that the assays are used at more than one cut-off in order to estimate the large number of parameters in the model. Only one study, by Land *et al.*,[274] met this requirement, using five different assays at between two and five different cut-offs. The study population, whose age was not reported, were recruited in an infertility clinic.

Having identified this study, our method further required information on the sensitivity and specificity of the assays used. These were the Biomérieux and Labsystems MIF assays and three peptide-based assays: the Medac pELISA, the Savyon ELISA, and the Labsystems ELISA (ELISA, enzyme-linked immunosorbent assay).

Three studies[265,268,269] were identified which provided evidence of sensitivity and specificity in these assays. We provide analyses of test resolution based on each of these sources (see *Appendix 16*). However, the estimates used in the analyses were those from Wills *et al.*,[269] for reasons explained below (see *Discussion*).

Analysis of serological assay performance

We characterise test performance as follows. The *resolution* (R) of each test T is a fixed property of the test, reflecting its intrinsic ability to distinguish between true positives and true negatives: high resolution implies greater discrimination. As the threshold titre at which a sample is declared 'positive' is moved from low titres to high titres, sensitivity is raised and specificity lowered. This traces out the receiver operator characteristic (ROC) curve,[277] which defines the sensitivity that is achieved for a given false-positive rate. Resolution for test T is defined as:

$$R_T = \text{logit}(sensivity_T) - \text{logit}(1 - specificity_T) \tag{70}$$

We use data on the sensitivity and specificity of a pgp-3 assay studies by Wills *et al.*,[269] and three peptide-based assays: the Medac pELISA, Savyon SeroCT ELISA and Labsystems ELISA. The mean R_T for these four assays was 4.1, with a SD of 0.8. The analyses on which this estimate is based are shown in *Appendix 16*, along with analyses of the performance of tests described by Morre *et al.*[265] and Narvanen *et al.*[268] To give an example of the relationship between resolution, sensitivity and specificity: if $R_T = 4.1$, and if we set sensitivity to 65%, then, from equation 70, specificity is 97%; with a sensitivity of 75%, we obtain specificity 95.3%.

Test sensitivity in individuals with an inflammatory reaction to Chlamydia trachomatis

Using a whole cell immunofluorescence (WIF) assay,[278] Akande *et al.*[270] reported that in women with antibody titres of > 64, and therefore likely to be previously exposed to CT, 84% (301/357) of those with tubal damage had titres of ≥ 256 compared with only 43% (83/193) of those with no signs of tubal damage (figures read from the published histogram). If the logit probability of a positive test result increases with log titre, then these results correspond to an increase in logit (sensitivity), and hence in test resolution of 1.96 (95% CrI 1.56 to 2.37). Similar results are found if a cut-off of 1 : 512 or 1 : 1024 is used. This provides an estimate of the difference in sensitivity, and hence test resolution, between those previously exposed to CT, with and without a CT-related inflammatory process: $\Delta_{Se} = \text{logit}(Se_{Max}) - \text{logit}(Se_{Previous}) = 1.96$.

Thus, based on results from Akande *et al.*,[270] we estimate that test sensitivity, and hence resolution, is 1.96 (95% CI 1.56 to 2.37) higher in this group than in women with previous CT but no inflammatory reaction. For example if resolution $R_T = 4.1$ in women who were previously infected with CT, but who have not experienced an inflammatory reaction, it would be $4.1 + 1.96 = 6.06$ in women with TFI caused by CT. This illustrates how the interpretation of sensitivity and specificity depend on the population being tested. For example, if we take 97% as a benchmark specificity, then the sensitivity in the group whose TFI was caused by CT would be 93%, but only 75% in CT-exposed women without an inflammatory reaction.

Analysis of a retrospective study of antibody prevalence in women with tubal factor infertility and control subjects

In a test T, given information on its sensitivity in a control group, $Se_{T,Previous}$, and a false-positive rate (Fp_T), the true prevalence in the control group, π_0, can be deduced from the observed prevalence $p_{T,Controls}$, via the equation:[55]

$$p_{T,Controls} = \pi_0 \cdot \text{Se}_{T.Previous} + (1-\pi_0) \cdot Fp_T \tag{71}$$

Here, the observed prevalence is conceived as a weighted average of true positives and false-positives, with π_0 as the weight.

The same reasoning can be applied to the prevalence in the TFI group. Their observed prevalence $p_{T,TFI}$ has three components. First, a proportion π_1 has been previously exposed to CT without it causing TFI. This group has a sensitivity $Se_{T,Previous}$. Second, there is a further proportion π_2 of TFI cases whose TFI was caused by CT. The sensitivity in this group is $Se_{T,TFI}$. Third, the remaining $1 - \pi_1 - \pi_2$ patients are true negatives, and so the prevalence of antibody in this group is the false-positive rate, Fp_T:

$$P_{T,TFI} = \pi_1 \cdot Se_{T,Previous} + \pi_2 \cdot Se_{T,TFI} + (1-\pi_1-\pi_2) \cdot Fp_T \tag{72}$$

We should expect that the prevalence of past CT infection in women with TFI *not* caused by CT is higher than in an otherwise comparable control group of fertile women. This is because most TFI is caused by STIs.[257] Therefore, women whose TFI is caused by *other* infections are more likely to have been exposed to CT. To capture this belief, we impose the constraint $\pi_1/(1-\pi_2) \geq \pi_0$ and rely on sensitivity analyses to assess its impact on results.

The Land *et al.* study[274] reports CT antibody prevalence in 51 laparoscopically verified TFI cases and 264 controls, on five different antibody tests, and at between two and five different cut-offs. The above model is fitted to this data set as described below.

In our base-case model, the difference Δ_{Se} in sensitivity between $Se_{T,Previous}$ and $Se_{T,TFI}$ is assumed to be the same for every assay. Alternative models are explored in the sensitivity analyses. We fit the model with R_T, Δ_{Se} and π_0 ranging over a wide 'grid' of fixed values, in a total $7 \times 3 \times 9 \times 5 = 945$ combinations, as follows:

- Test resolution R_T associated with each test T, is assumed to come from a normal distribution with means fixed at 0.5, 1.5, 2.5, 3.5, 4.5, 5.5 and 6.5, and SDs fixed at 0.5, 0.8 or 1.1. As reported above, the mean and SD in the Wills *et al.* study[269] (above) were 4.1 and 0.8, respectively. To give an idea of what this range in R_T would mean: with sensitivity 60%, $R_T = 0.5$ implies 52.4% specificity, whereas $R_T = 6.5$ implies 99.8% specificity.
- We fixed Δ_{Se} values at 0, 1, 2, 3, 4, 5, 6, 7 and 8. Based on Akande *et al.*,[270] $\Delta_{Se} = 1.96$. Values above 3 or 4 imply sensitivity in the region of 100% at typically accepted levels of specificity.
- The prevalence of previous CT exposure, π_0, i.e. the cumulative incidence, in the control group, was set to 25%, 30%, 35%, 40% and 45%.

The parameter estimates and goodness of fit for each model were recorded.

Statistical estimation

Appendix 16 details the prior distributions, and describes methods for taking account of the uncertainty regarding the true status of serological samples. *Appendix 17* gives the WinBUGS code and data sets used to compute the reported results.

Results

Goodness of fit of the data in the Land *et al.* study,[274] based on residual deviance, is plotted against mean resolution, with different values of Δ_{se} in *Figure 27*. In this graph, π_0 has been held fixed at 0.25, and SD at 0.8. Based on the size of the data set, a model that fits well would be expected to have a residual deviance of 34. Residual deviance of > 40 suggests a conflict between data and model. If we adopt a residual deviance of 40 as a maximum, the evidence in the Land *et al.* data suggests that values of R_T of > 3.5 and values of Δ_{se} of < 2 are implausible, although $\Delta_{se} = 2$ is the best fitting. The poor fit obtained when $\Delta_{se} = 0$ or 1 provides strong support for our interpretation of Akande *et al.*'s results.[270]

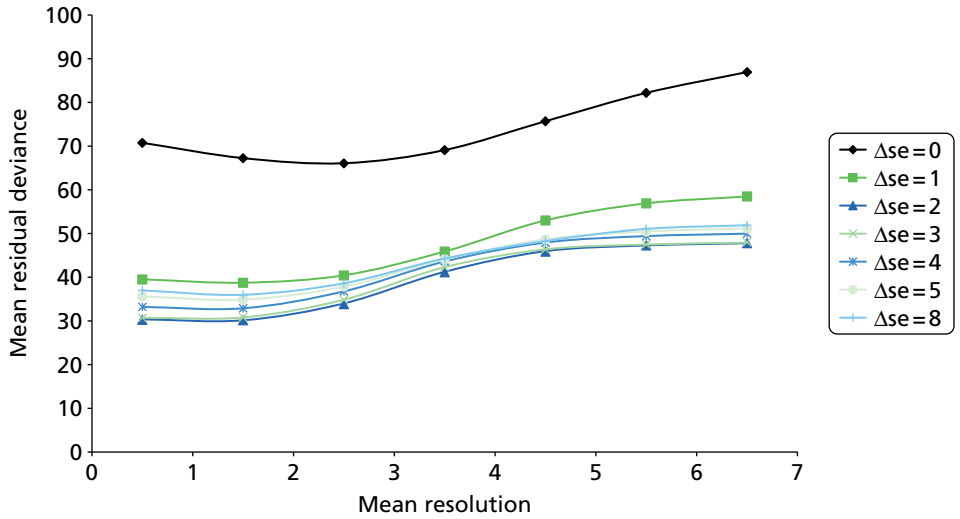

FIGURE 27 Mean residual deviance of models fitted to the Land data plotted against mean resolution of the serological assays, showing the impact of assumptions on Δ_{se}, the difference in test resolution between those exposed only to CT and those suffering an inflammatory reaction. From Price MJ, Ades AE, Welton NJ, Macleod J, Turner K, Simms I, *et al.* How much tubal factor infertility is caused by chlamydia? Estimation based on serological evidence adjusted for sensitivity and specificity. *Sex Transm Dis* 2012;**39**:608–13.[275]

The estimated mean PEF (π_2) is plotted against baseline prevalence of past infection in the control arm (π_0) (*Figure 28*). This is restricted to models with SD = 0.8 and mean residual deviance of less than 40. Also excluded are models with a mean resolution below 2, which is unrealistically low, given Wills *et al.*'s findings[269] (see *Appendix 16*). A reasonable maximum Δ_{se} was set at 3, and this restriction was also applied.

To take account of the uncertainty in choice of model, we have applied an averaging procedure to the *Figure 28* models. This uses the smoothed residual deviance to calculate a weighted parameter estimate.[279] This generates a central estimate of π_2, the aetiological fraction, of 0.45, with CrI (0.28 to 0.62). Finally, the observed prevalence in TFI cases and controls reported in the Land study is compared with the predictions from the best-fitting model: (resolution = 2.5, $\Delta_{se} = 2$, and $\pi_0 = 25\%$, with residual deviance of 34), in which the aetiological fraction was estimated to be 50% (95% CrI 37% to 60%) (*Figure 29*).

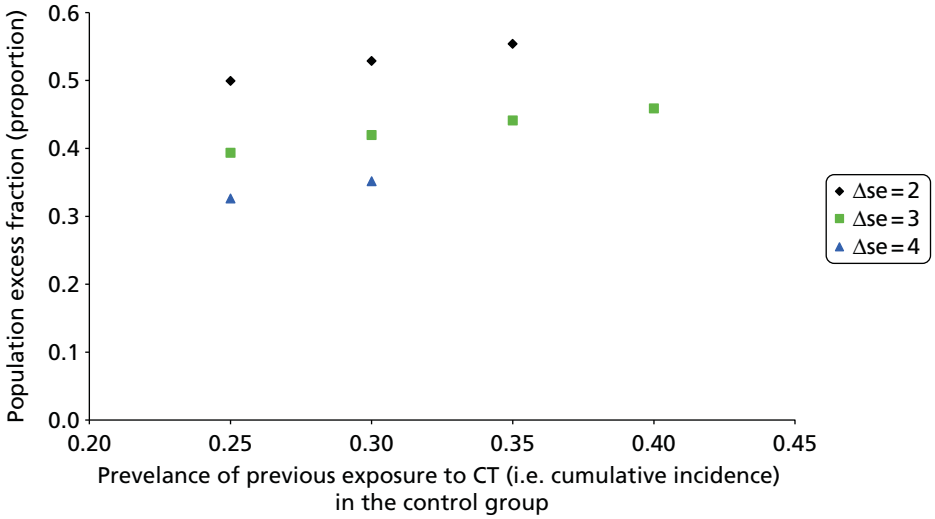

FIGURE 28 Population Excess Fraction, given various assumptions about the difference in resolution between those previously exposed to CT and those with an inflammatory reaction to it, and assumptions about the prevalence of previous exposure to CT. From Price MJ, Ades AE, Welton NJ, Macleod J, Turner K, Simms I, *et al.* How much tubal factor infertility is caused by chlamydia? Estimation based on serological evidence adjusted for sensitivity and specificity. *Sex Transm Dis* 2012;**39**:608–13.[275]

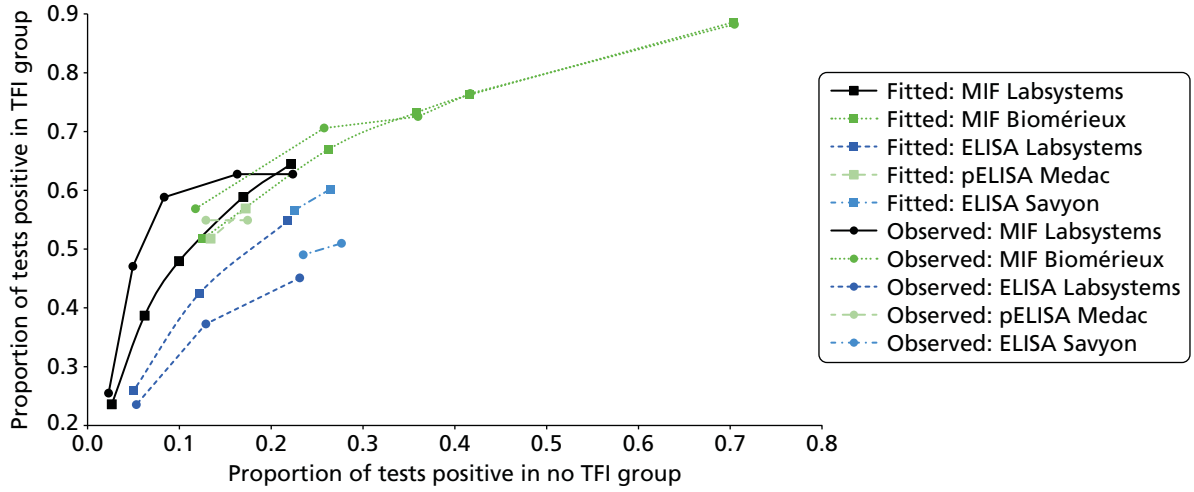

FIGURE 29 Comparison of observed prevalence of antibody +ves in the Land study with the predictions from the best fitting model. From Price MJ, Ades AE, Welton NJ, Macleod J, Turner K, Simms I, *et al.* How much tubal factor infertility is caused by chlamydia? Estimation based on serological evidence adjusted for sensitivity and specificity. *Sex Transm Dis* 2012;**39**:608–13.[275]

Sensitivity analyses

Altering the SD over the range 0.5 or 1.1 does not have a material impact on results (not shown). Increasing π_0 to 45% increased the residual deviance by around 15% at all points on *Figure 27*, but did not change the overall pattern of results.

The model identified as best fitting in *Figure 27* shows a systematic lack of fit. Fitted values for the control group in this model are very close to the data, as can be seen from their proximity on the horizontal axis (see *Figure 29*). For the TFI cases, the model *under*predicts some of the data where the detection rates are highest (particularly the MIF tests), but *over*predicts the data where the detection rates are low (particularly the Labsystems ELISA and the Savyon ELISA). A similar pattern is found for all sets of parameter values.

These results point to an apparent anomaly in the Land data: some tests detect more positives in the control group than others yet identify *fewer* positives in the case group. One explanation is that assays differ in which CT serovars they are sensitive to. Thus, the detection threshold can be lowered, raising the false-positive rate without detecting any more cases. The low correlation sometimes observed between titres on different tests may be a reflection of this.[280] As a second sensitivity analysis, we performed a post-hoc analysis in which all parameters were held at the same values, but the Labsystems and Savyon ELISAs were able to detect only a proportion of CT cases detected by other tests. This model fitted well (residual deviance of 25) and removed the systematic lack of fit (*Figure 30*). The Labsystems detected an estimated 60% (95% CrI 42% to 82%) and the Savyon 69% (95% CrI 51% to 88%) of CT cases. The attributable fraction was 58% (95% CrI 47% to 67%), somewhat higher than the model in the main text, but suggesting that results are not overly sensitive to this aspect of the model assumptions.

A third sensitivity analysis explored the assumption that Δ_{Se} was the same in all tests. Alternative models were explored in which the mean R_T is set to 2.5 or 3.5, with a between-test SD of 0.8, as in the main analyses, but Δ_{Se} was allowed to vary between tests with mean values set at 2 or 3 and with between-test SD set at 1 or 1.5. The eight models all fitted well (residual deviance between 27 and 36). Estimates of the aetiological fraction ranged from 44% (95% CrI 31% to 54%) to 51% (95% CrI 38% to 62%).

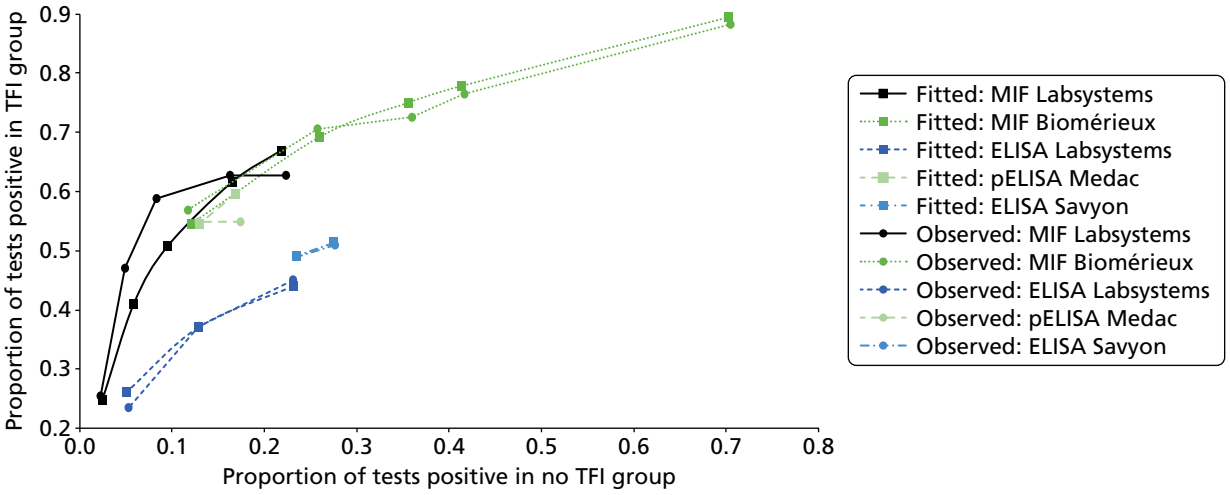

FIGURE 30 Sensitivity analysis: observed and posterior mean fitted values for each data point in the Land study when the ELISA tests are assumed to only detect a proportion of chlamydia serovars. From Price MJ, Ades AE, Welton NJ, Macleod J, Turner K, Simms I, *et al.* How much tubal factor infertility is caused by chlamydia? Estimation based on serological evidence adjusted for sensitivity and specificity. *Sex Transm Dis* 2012;**39**:608–13.[275]

A fourth sensitivity analysis was run to assess the impact of dropping the constraint $\pi_1/(1-\pi_2) \geq \pi_0$, which builds into the models the assumption that women with TFI whose TFI is *not* caused by CT are likely to have a higher probability of prior exposure to CT than control subjects. Twelve models were evaluated ($R_T = 2.5, 3.5$; $\Delta_{Se} = 2,3$; $\pi_0 = 0.25, 0.35, 0.45$). The aetiological fraction was between 0 and 6.3 percentage points higher. In the four models that conformed to our criterion for goodness of fit, i.e. mean residual deviance of < 40, the increase was between 0 and 4.3 percentage points.

Discussion

In this chapter we have derived estimates of the performance of contemporary serological assays for CT in different populations, based on literature. This literature suggests that CT antibody levels, and with it test sensitivity, depend on the population tested. Specifically, antibody levels – and thus sensitivity – are higher in women with TFI caused by CT than in women whose TFI was caused by another organism, or in women with a previous CT infection and no TFI. The model of serology induced from this evidences was used to obtain estimates of the PEF, the proportion of TFI that is caused by CT from a published case–control study comparing antibody prevalence in TFI cases compared with non-TFI controls.

We estimated that 45% (95% CrI 28% to 62%) of TFI cases were caused by CT. Although the CrIs are wide, consistent results were obtained from a range of models. It was also possible to demonstrate a fair degree of insensitivity to reasonable variation in modelling assumptions.

Our approach contrasts with the standard method of estimating PEF from ORs. Rather than control for the positive confounding between TFI and CT that is not causally related, which would be difficult to implement in this data set, we model the PEF directly based on assumptions about the titre distributions in cases and controls. If we apply the standard formula to the Land data as it stands, we obtain PEF of between 28% (95% CrI 13% to 44%) and 60% (95% CrI 37% to 78%).

The study gives us some fresh insights into the interpretation of sensitivity and specificity as typically reported, and suggests that contemporary CT antibody assays may be less sensitive than has been reported. The test resolution of 4.1 derived from the Wills *et al.* study[269] would imply 71% sensitivity at 96% specificity. However, judging from our results from the Land data set, it appears that resolution is, at best, only 3.5. This would imply, at best, 58% sensitivity at 96% specificity. Although disappointingly low, in a population where most antibody positive women are likely to have already cleared their CT infection, this is probably a reasonable estimate. Sensitivity is normally defined in currently infected individuals, and antibody levels are known to decline with time since infection.[281] Previous experience with a pgp3 Indirect ELISA assay, suggested that sensitivity was 89% (95% CrI 75% to 98%) within 6 months of infection, but no more than 64% (95% CrI 51% to 77%) 2 years after the last known infection. This is equivalent to a 1.5-point drop in test resolution after infection has cleared.

Two further studies[265,268] of test performance in peptide-based assays are re-analysed in *Appendix 16*. In one study,[265] a discrepancy analysis was used to examine the same three peptide-based assays and a MIF test. We estimated an average test resolution of 5.9, possibly due to the well-recognised tendency of discrepancy analysis to exaggerate sensitivity at a given specificity.[266,267] An even higher estimate of resolution, 6.29 (95% CI 4.84 to 8.19), was obtained from a second study of an EIA peptide assay,[268] which was sponsored by its manufacturers. In this study true positives were identified using CT culture. As this is much less sensitive than the NAAT tests used by Wills *et al.*[269] we can expect that the manufacturer study preferentially selected patients with more severe infections, and therefore higher antibody levels.

We observed above that raised antibody levels have often been reported in women suffering from complications of CT. Our analysis of the Land data decisively rejects the hypothesis that sensitivity in women with CT-related TFI is the same as in women previous CT infection and no complications. The analysis further suggested a 2-point difference in resolution between these groups, which is exactly consistent with the findings from the Akande *et al.* study.[270]

We must conclude that most published estimates of CT antibody test performance lack face validity, as it depends wholly on the criteria and method for patient selection. Among the critical factors are: current infection versus previous infection; NAAT versus culture to identify true positives; methodology to identify true negatives; and inclusion versus exclusion of women with an inflammatory response to CT. This tells us that, although the development of assays with superior performance must be welcome, methods for *characterising* test performance should perhaps be given even higher priority.

Limitations of the analysis

The study has relied on a series of assumptions about parameter values. Although we have done as much as possible to justify these values and base them on published literature, further systematic study is needed. The final estimate of the PEF emerging from our study (45%, 95% CrI 28% to 62%) has wide credible limits, reflecting, in part, the uncertainty about parameter values and modelling assumptions. Even so, we suspect that other assumptions and other models could be found to fit these data, so the uncertainty in the estimate may yet be understated.

A second set of limitations relate to the extent to which the results generalise to the UK situation. As noted in earlier chapters, estimates of PEF are specific to time and place, and it is not known how close the spectrum of infections leading to salpingitis in the Netherlands and the UK might be. Given the decline in the CT PEF estimates in PID (see *Chapter 8*), in projecting the results to the UK we must certainly take into account the age of the women recruited into the Land study. The authors do not report an age range, but as the study was based in an infertility clinic, we can infer that, on average, women would be older than typical pregnant women, and the majority would be aged > 24 years.

A further concern about generalisability relates to the specific selection procedures in the Land study. The authors reported that MIF titres were assessed at the first visit and were used to assess TFI risk. Those with titres of > 64 had laparoscopy to establish TFI status, whereas others received laparoscopy only after a positive result on hysterosalpingography. This may have induced a reduced probability of recruiting TFI cases with low titres. Similarly, the laparoscopic criteria for a diagnosis of TFI was selected on the basis that it had been shown previously to 'reflect chlamydia-related tubal disease most accurately'. The authors noted that selection and verification biases were difficult to avoid and that likelihood ratios would be overestimated by the procedures followed. It is not clear how the extent of the upwards bias in PEF estimates can be assessed.

Looking to the future, the study suggests that a greater role for direct modelling of the distributions of titres in TFI or EP cases compared with carefully chosen controls. This could be invaluable both in aetiological studies of CT and its sequelae, and in surveillance of CT for purposes of prevention and control. Serology could help assess the impact of public health interventions, including the NCSP, to reduce the population burden of genital CT.[264] The possibility that serology could improve our estimates of the burden of CT, and contribute to understanding its natural history,[38,98] will be taken up later in the research recommendations.

Summary of assumptions and findings

Summary of assumptions and limitations

1. TFI caused by CT is associated with higher serum antibody levels than previous CT infection not leading to TFI.
2. Sensitivity of CT antibody tests depends on the stage of CT natural history.
3. It is not known how closely the proportion of TFI attributable to CT (PEF) in the Netherlands approximates the PEF in the UK.

Summary of findings

1. High CT titres are associated with TFI.
2. CT titres can be used to estimate the proportion of TFI attributable to CT.
3. The estimated PEF from the Land study is 45% (95% CrI 28% to 62%).

This estimate relates primarily to women aged > 24 years, and, in view of the age gradients reported in *Chapters 7* and *10*, would be likely to be an *underestima*te of the PEF for the female population aged 16–24 years.

Chapter 12 Summary and conclusions

Objectives

To:

1. identify and, if necessary, reconcile inconsistencies in the results presented in previous chapters
2. summarise the additional risks of poor reproductive outcomes following from CT infection in individual women
3. summarise the public health significance of the research
4. identify remaining areas of uncertainty and make research recommendations to address them.

Introduction

In this final chapter we review the results of previous chapters, and ask whether or not, when everything is put together, this programme of work has generated a consistent picture of the clinical epidemiology of CT and its *sequelae* in the UK, as it was in 2002. We then consider what further assumptions are needed to reconcile inconsistencies that remain.

Having proposed what we regard as the most plausible overall model, consistent with all sources of evidence, we then set out the implications for individual women and for public health. We include some brief comments summarising the contribution of this report to the study of the clinical and population epidemiology of chlamydia, and the role of the MPES methodology. We conclude by proposing research recommendations, based on our judgement of the key areas of uncertainty and the research designs that have the best chance of decreasing the uncertainty in policy-sensitive parameters.

Summary of findings from previous chapters

Chlamydia trachomatis and pelvic inflammatory disease

Chapters 4–7 provided a comprehensive account of the CT incidence, prevalence and duration, the proportion of CT that is symptomatic, the incidence of diagnosed 'probable/definite' PID and total PID (diagnosed and undiagnosed), the CT-to-PID progression risk, and the proportion of PID attributable to CT.

Summary

The main results were as follows (see chapter summaries for more detailed quantitative results):

1. The extreme heterogeneity of results in studies of CT duration could be effectively explained by (1) distinguishing screening studies of prevalent infection from clinic studies of incident infection and (2) assuming that a proportion of infections are 'passive', which accords with existing biological understanding of infection (see *Chapter 2*).
2. The estimate of duration is consistent with independent sources of evidence on age-specific CT incidence in the UK, age-specific CT prevalence in the UK, and the proportion of infections that are asymptomatic.
3. We derived an estimate of the CT-to-PID progression risk from CT screening trials. This was based on a Markov model that allowed for CT clearance as a 'competing risk'. Trials of different designs gave consistent results, with a pooled estimate of the CT-to-clinical PID risk of 14.8% (95% CrI 4.8% to 24.8%).

4. We interpret this progression risk as representative of a level of ascertainment of PID that can be expected to be seen in prospective studies of women at risk of PID. The risk of progression to *all* PID, including symptomatic and asymptomatic, is 17.1% (95% CrI 5.6% to 29%).
5. Two independent sources of evidence on the incidence of clinical PID were derived: one from the control arm of the POPI trial[19] taking into account the degree of ascertainment expected in prospective studies, the other from routine GPRD, HES and KC-60 data, taking account of overlap, and of independent data on the proportion diagnosed. The two sources were in agreement, and a pooled incidence of PID is 2.5% per year in women aged 16–24 years, and 1.8% per year in women age 16–44 years.
6. If age-specific CT incidence, age-specific PID incidence and CT-to-clinical PID progression risk are put together, assuming progression risk is independent of age, we predict that the proportion of clinical PID attributable to CT is 35% in the 16–24 years age group, and 20% overall (age 16–44 years).
7. These predictions of the PEF are in agreement with (a) estimates based on retrospective studies of CT infection in PID cases and controls, corrected for under-ascertainment of infection in cases, and (b) estimates based directly on the POPI trial corrected for independent testing in the control group. However, there is considerable uncertainty in all these estimates.

In our analysis of the PEF there was an element of post-hoc reasoning: this arose from conflicts between our initial estimates and results of later chapters (see *Chapter 7* discussion).

Clinical pelvic inflammatory disease, salpingitis, ectopic pregnancy and tubal factor infertility

Chapter 8 developed estimates of the proportion of women who have experienced PID, salpingitis and subsequent episodes of PID in the UK. In *Chapters 9* and *10*, the risks of EP and TFI in women with salpingitis that were observed in the Lund study were then applied to these risk factor distributions to predict the risk of EP per pregnancy and the prevalence of TFI in the UK. *Chapter 11* reports an estimate of the PEF (CT←TFI) based on analyses of serological data. More detailed results appear in the chapter summaries.

Summary

1. The proportion of 'probable/definite' PID in which salpingitis can be confirmed on laparoscopy is estimated to be 43%.
2. An estimated 16.1% of 35- to 44-year-old women have experienced at least one episode of salpingitis.
3. Based on HES data, the rate of ectopic pregnancies as a proportion of pregnancies rises with age and averages 1.1% in women aged 16–44 years.
4. Based on UK fertility surveys, 1.1% of women aged 44 years have TFI (including both primary and secondary infertility).
5. Of EPs, 27% are due to salpingitis, based on French case–control studies.
6. Of EPs, 4.9% are due to CT.
7. Given our core assumptions, the observed EP incidence in the UK is compatible with the Lund study if it is assumed that salpingitis in those whose PID is not diagnosed in hospital has the same risk of EP as was observed in those with 'mild' salpingitis, and that undiagnosed salpingitis carried no risk of EP.
8. Given our core assumptions, the observed TFI prevalence in the UK is compatible with the Lund study if it is assumed that the risk of TFI following salpingitis is the same as, or slightly less than, the average risk observed in the Lund study, in which the salpingitis is associated with diagnosed or undiagnosed PID.
9. Based on our analysis of the proportion of salpingitis due to CT and the relation between salpingitis and TFI, 29% of TFI is due to CT.
10. In an analysis of retrospective serological data collected from women attending an infertility clinic in the Netherlands, 45% (95% CrI 28% to 62%) of TFI was due to CT.

Conflicts in the overall evidence and their resolution

Two clear points of conflict emerge in the evidence presented in the report:

1. The models that correctly predict the risk of TFI following salpingitis seriously *over*-predict the risk of EP following salpingitis.
2. The two estimates of the proportion of TFI due to CT are inconsistent, one based on our prospective analysis of salpingitis and TFI, the other on retrospective serological data.

Ectopic pregnancy and tubal factor infertility

It is important to note that, to the extent that we are correct in assuming that salpingitis-related reproductive damage represents a single common pathway to EP and TFI, the conflict between prediction for EP incidence per pregnancy and TFI prevalence is largely *independent* of the way we have analysed the Lund data on EP and TFI, and also of our analyses of the incidence of PID, the cumulative incidence of PID, the relation between clinical PID and salpingitis, and the role of undiagnosed PID. Whatever the precise truth about these factors, they should be operating the same way in EP as in TFI.

One major difference between EP and TFI in this context is that salpingitis is not the only cause of EP, but it is virtually the only cause of TFI.[35] As a result, the Lund study findings for EP are more vulnerable to confounding, whereas those on TFI are not. One well-recognised confounding factor is smoking, which is likely to have been more prevalent in women with more exposure to STIs. A second possible confounder is IUD use. The role of the earlier IUDs in the aetiology of EP is not entirely clear. Chavkin[282] reviews mechanisms that could generate a positive confounding effect. First, it is likely that IUDs are highly effective in preventing uterine implantation, but much less effective at preventing tubal implantation, so that women with IUD are selectively more likely to experience an EP. However, this does not explain why no increased risk of EP was observed in the Lund study in women with IUDs but without salpingitis.[110] A second possibility is that the older-generation non-medical IUDs were associated with some form of non-bacterial salpingitis that raises the risk of EP in the absence of any infective mechanism. This seems to be ruled out by the observation that women with no visible infection on laparoscopy did not have elevated risks of EP or TFI.[283]

A third proposal was that the presence of an IUD could have raised the risk of PID-associated EP. On this model, IUD is not so much a confounder as a risk-modifier. Based on changes in salpingitis exposure and changes in IUD exposure, the Lund investigators themselves attributed a high proportion of the observed doubling in EP rates during the course of their study to an increase in IUD use.[110] It is likely, indeed, that the Lund study could even have underestimated the role of IUDs, because the pain and discomfort associated with PID and IUDs can be similar, and many IUDs could have been removed prior to entry to the study.

It is unlikely that further study will resolve this issue, because the IUDs used during that period have been replaced by medicated IUDs, which are much less likely to cause problems.[77,114] We can, therefore, only offer the suggestion that the EP progression rates in the Lund study may be greatly overestimated through the presence of confounders, such as smoking, or effect modifiers, such IUD use. As there seems to be no way of estimating the extent of this bias independently, in coming to final conclusions we will effectively discount the prospective Lund data on EP risk entirely, and accept salpingitis-related EP rates based on the PEF from the retrospective French studies applied to the observed HES data, which give between 0.24% and 0.37% of all conceptions as salpingitis-related EPs.

We note again that this estimate is based on a post-hoc re-interpretation of the evidence, after examination of other results, and therefore requires independent confirmation.

Causal role of Chlamydia trachomatis in tubal factor infertility

The applicability of the PEF from the serological data from the Netherlands to the UK must be considered doubtful. This is not so much because of the uncertain relevance to the UK of evidence from the Netherlands, but, more seriously, because the study authors explicitly warn of the high risk of verification and selection biases in the study, which would artificially raise the PEF (see discussion in *Chapter 11*).

We therefore proceed with the lower estimates based on the *Chapter 10* analyses – 29% in women aged 16–44 years – but note once again a post-hoc element to this decision.

Final set of coherent estimates

A set of coherent estimates for the parameters covered in *Chapters 5–7* was provided in those chapters. The retrospective estimates on EP (see *Chapter 9*) and the prospective estimates on TFI (see *Chapter 10*) are also technically coherent with estimates from earlier chapters, as they are based on estimates of salpingitis prevalence from *Chapter 8*, which are based on forward simulation from the estimates in *Chapters 5–7*, and there is no feedback of the evidence on EP and TFI back on to these earlier parameters.

Future investigators citing or using the marginal results directly from the tables included throughout the report are accessing a fully coherent set of estimates. However, these summaries are insufficient to characterise the full joint distribution of the parameters, as they do not express the correlations between parameters. Investigators using our results in ways that take uncertainty into account should therefore obtain the joint parameter distributions by running the WinBUGs code in the appendices.

Significance for patient, clinicians and public health

Summary outcomes and contribution of Chlamydia trachomatis *in the population*

Table 42 gives an impression of the numbers of PID, EP and TFI outcomes that occur in England each year, and the numbers attributable to CT, based on our estimates of PEF in previous chapters. Based on our estimate that the incidence of CT in women aged 16–44 is 2.1 per 100 person-years (see *Chapter 7*), we can use the information in *Table 42* on the numbers of sequelae attributable to CT, to calculate the risk of sequelae following CT infection. Every 1000 CT infections on average give rise to approximately 171 PID episodes, 73 salpingitis episodes, 2.0 EPs, and 5.1 women with TFI by age 44.

Public health implications

We have estimated in *Chapter 7* that that only 12.6% of PID is asymptomatic, although a much higher proportion is undiagnosed in practice. These findings require confirmation from large contemporary studies, but, if accepted, they have the important public health implication that, *in women who are aware of the risk of PID and aware of the symptoms,* few PID episodes will be missed. This suggests that much could be gained from improved public understanding of signs and symptoms that should lead women to seek early medical attention to avoid the risk of reproductive damage.

Our findings do not suggest any changes to current guidelines on diagnosis and management of PID, which advise early treatment with broad-spectrum antibiotics. However, there is a degree of ambiguity and inconsistency in current UK guidance, which focuses on younger women and possibly overemphasises the role of STIs (see *Chapter 2*). It is relevant, here, that our analyses of PID incidence in *Chapter 7* suggest that about 63% of PID occurs in women aged > 25 years.

Although the trial evidence demonstrates that treating prevalent asymptomatic CT infection is effective in preventing clinical PID, we estimated that, at best, a woman aged 16–24 years undergoing screening at annual intervals would on average prevent only 61% of CT-related PID. Given our analysis of the

TABLE 42 Summary findings: frequency (95% CrI) of new incident PID, EPs and TFIs caused by CT in England per year

	PID			EP			TFI		
Age (years)	Total per year (A)	Percentage due to CT (B)	CT caused per year (C = A × B)	Total per year (A)	Percentage due to CT (B)	CT caused per year (C = A × B)	Total per year (A)[a]	Percentage due to CT (B)	CT caused per year (C = A × B)
16–19	25,490 (18,510 to 34,880)	53.50 (16.58 to 100)	13,330 (4093 to 25,660)	291	3.05 (0.79 to 7.08)	9 (2 to 21)	84 (39 to 164)	52.7 (16.1 to 100)	44 (11 to 110)
20–24	42,330 (30,330 to 58,460)	24.32 (7.27 to 46.67)	10,050 (3116 to 18,720)	1154	5.49 (1.52 to 12.32)	63 (18 to 142)	287 (143 to 527)	38.1 (11.9 to 74.5)	99 (26 to 237)
25–34	67,810 (45,470 to 97,960)	10.25 (2.92 to 21.18)	6 711 (2026 to 12,960)	4340	5.33 (1.55 to 14.44)	231 (67 to 627)	1 358 (802 to 2 191)	25.8 (8.1 to 50.5)	328 (92 to 732)
35–44	47,700 (29,760 to 71,870)	11.47 (3.01 to 25.78)	5 211 (1495 to 10,720)	1840	4.67 (1.30 to 9.99)	86 (24 to 184)	2164 (1822 to 2 757)	20.0 (5.9 to 40.0)	345 (99 to 689)
16–44	183,300 (131,000 to 253,800)	19.74 (5.93 to 38.06)	35 300 (11,100 to 64,560)	7625	4.92 (1.40 to 10.59)	375 (107 to 807)	3893 (2806 to 5639)	28.7 (8.9 to 56.3)	817 (227 to 1768)

a Based on Model 1 in *Chapter 10*.
Total events in England (A), proportion caused by CT (B) and frequency of CT-caused incident PID, EP, TFI (C = A × B).

proportion of PID that can be attributed to CT, it appears that annual screening would directly prevent no more than 22% (95% CrI 7% to 43%) of all-cause PID in this age range.

Although our estimates of the risks of PID and salpingitis are high compared with previous estimates, this assessment of the benefits of annual testing to individual women is relatively pessimistic. Taken together, the findings suggest that greater benefit to individual women could be achieved by increased attention to incident infection, in other words a greater focus on women at an increased risk of recent exposure, as is the intention of current policy that advises screening after partner change, and re-screening 3 months after a diagnosis (acknowledging a high risk of re-infection). This could also suggest testing women following unprotected sex with a high-risk partner. Such strategies should have a higher yield, in terms of the number of infections detected per test carried out, and should also increase the chance that treatment would prevent the development of PID. This also fits well with an integrated public health approach to all STIs (chlamydia, gonorrhoea, human immunodeficiency virus (HIV), syphilis, genital herpes, genital warts and trichomonas) in the general population, with an emphasis both on prevention and on the early detection of infection and treatment, and effective management strategies to reduce onwards transmission, for example through rapid partner notification.[284]

The study has also clarified both the incidence of PID and the risks PID poses to reproductive health. These may be greater than has been appreciated. We have estimated a relatively high incidence of PID in women aged 16–44 years, 1.8% per year, of which 66% undiagnosed. But we have also shown that it is difficult to explain the UK prevalence of TFI, unless it is assumed that the risk of salpingitis is the similar in diagnosed and undiagnosed PID and that salpingitis, whether diagnosed or not, carries a similar risk of TFI – at, or close to – the level of risk reported in the Lund study.

It has been suggested, based on a systematic review of prospective studies, that there is no evidence for a causal link between CT and TFI.[145] Our estimate is that 29% of TFI is due to CT. Although the CrIs on this estimate are wide (95% CrI 9% to 56%), it derives from our estimates of the proportion of PID due to CT. In addition, although it is difficult to obtain reliable quantitative estimates of the PEF from serological and microbiological studies, an extensive body of literature, some of which is cited in *Chapter 2*, as well as our analyses in *Chapter 11*, confirms that CT is strongly implicated in TFI.

Effective commissioning of interventions and services is key to improving outcomes. Most sexual health services will be commissioned by local authorities, but Clinical Commissioning Groups (CCGs) and the NHS Commissioning Board (NHS CB) will also have a role.

Priorities for future research

It is possible to envisage more randomised studies of screening, perhaps involving different populations, infection ascertainment, treatment or partner-tracing methods. However, there is little realistic foreseeable prospect of interpretable comparative studies following CT infection to longer-term outcomes. The obligation to treat CT infection makes non-randomised designs, including linked register studies, virtually uninterpretable. In addition, the extended follow-up required to assess reproductive outcomes inevitably leads to dilution of the effects, as a proportion of women in *both* comparison groups will seek testing and treatment independently, acquire subsequent CT infections, and experience episodes of PID.

In view of the difficulties associated with prospective studies, we outline, below, a series of relatively easy to conduct field studies based on routine health service activity. Further dynamic modelling using a MPES approach is also required to assess the cost-effectiveness of the NCSP as currently configured, extensions to it, and other chlamydia preventive measures.

Serological case–control studies and the age profiles of *Chlamydia trachomatis*, pelvic inflammatory disease, ectopic pregnancy and tubal factor infertility

Serological case–control studies were popular over 20 years ago, when they were used to throw light on the aetiology of these conditions; there was also an interest in using CT serology in treatment decisions. On a few occasions they have been used to estimate a PEF.

These studies fell out of favour because of the poor sensitivity and specificity of the assays, and also because these properties were difficult to characterise. However, better performing assays are now available, and there is better information on their performance characteristics. More recent assays have high specificity,[269] and the relation between the measured 'sensitivity' and the time since the infection is better understood.[281]

A number of investigators have reported that exceptionally high CT antibody titres can be found in a proportion of women with PID, EP or TFI,[88,270,271,273,285] with many studies reporting a clearly bimodal distribution of positive titres. These high titres have been attributed to an upper genital tract inflammatory response.[61,113] Using finite mixture models,[286] it would be possible to draw conclusions about the proportion of PID, EP and TFI cases that can be associated with these high CT titres. Given suitable control samples, and assuming that the high titres are indicative of causation, this form of analysis could provide direct estimates of PEF for CT.

Such studies could also be used to explore other assumptions made in this report. First, they could confirm or disconfirm whether the age profile of the PEFs accord with the predictions in this report. Variation with age in the proportion of PID, salpingitis, EP and TFI that is due to CT, is of considerable scientific interest and public health importance. Second, information could be collected on sexual and reproductive history, referral pathways, and, particularly, on previous clinical PID and previously unreported related symptoms. This could improve our quantification of all-cause PID through routine data, and the level of completeness of ascertainment that is achieved in trials. The study on which our estimates are based had only 36 TFI cases.[103] Possible risk factors for salpingitis, such as IUD use, could also be re-examined. Finally, microbiological evidence,[287] as well as our own analyses of the proportion of PID attributable to CT, suggests that the majority of PID is *not* due to CT. Possibly, these non-CT causes of PID become progressively more important with age. If this can be confirmed in microbiological studies, this would add further to our understanding of the aetiology of salpingitis (see *Chapter 2*).

A relatively simple and inexpensive series of case–control studies could, therefore, provide a great deal of information about the incidence of undiagnosed PID, the role of CT in diagnosed and undiagnosed PID, the role of CT in EP and TFI, and how that role depends on age. If this is successful, a set of sentinel clinics could be established for each of these surveys, providing a stable platform for ongoing surveillance of trends in PEF of CT, to be considered alongside surveillance of trends in overall EP and TFI rates.

Dynamic modelling

This report has focused on the risks of CT in downstream PID, salpingitis, EP and TFI. It has also touched on the role of screening in prevention of PID. The focus has therefore been on the reproductive risks to the individual women exposed to CT infection and the benefits of screening. The other objective of screening is to lower CT prevalence, and therefore incidence, to reduce the number of CT infections and downstream complications in the wider population.

Further dynamic modelling is now required, incorporating the new information on natural history that has been reported in this study, in order to fully assess the potential role for screening, as implemented in the NSCP, or in other ways. There is a wide consensus that the cost-effectiveness of screening is strongly determined by the risk of downstream illness and reproductive damage,[1,2] and this is relevant to both the individual and the population motivations for screening. The new information contained in this report is therefore critical to new modelling exercises in the UK.

At the same time, it is evident that the methodology of dynamic modelling needs itself to be developed further using MPES. This will enable more evidence to be incorporated into the models, and will allow the consistency of the evidence sources to be evaluated.

A distinction is often made between 'mathematical' models, which make predictions about long-term changes in incidence and prevalence, based on estimates previously derived from 'statistical' models of data,[288] suggesting a two-stage approach. However, in a multi-parameter evidence synthesis, the same mathematical models are estimated from the available data in a single step, but in a way that allows proper statistical inference and uncertainty propagation. In MPES there is no artificial distinction between evidence sources that are seen as model 'outputs' and those that are considered model 'inputs' in the standard mathematical modelling literature, or between data used for model 'calibration' and data used to estimate 'predictions'. These distinctions inevitably lead to failure to incorporate all the evidence.

Fundamentally, the MPES framework ensures that the (mathematical) model that generates the predictions is exactly the same model that generates the evidence. Once a certain level of complexity is reached, this cannot be achieved by a two-stage approach (see *Chapter 6*).

Although there are technical difficulties in applying the MPES methodology in infectious disease dynamic modelling, there are already case studies that establish its feasibility.[289,290] However, combination of infectious disease transmission dynamics with sexual network dynamics raises computational challenges, and the development of other methods of Bayesian computation besides, or in addition to, MCMC has become an active area of research.[291]

Contribution of this report

Many of the estimates in this report are similar to estimates that have appeared previously in the literature. Others, such as the estimates of the proportion of the population exposed to salpingitis, the consequent risks of EP and TFI, and the proportion of these outcomes that can be attributed to CT, have been based on new methods and new types of evidence, or have not been attempted at all. Of greater significance than the individual estimates, however, is the fact that, taken together, they form a coherent and internally consistent set of estimates which are at the same time consistent with an extensive and multifaceted body of evidence, including routinely collected statistics on PID, EP and TFI.

The parameters estimates are, in many cases, imprecise, but they have been estimated with as much precision as the evidence allows. Unlike previous estimates, the evidence base for each parameter has been explicitly delineated and, wherever possible, cross-validated against other sources of evidence. In the course of the project, inconsistencies between sources of evidence have been revealed, making it inevitable that evidence must be re-interpreted in order for inconsistencies to be reconciled. Where we have resorted to post-hoc reasoning of this sort, this has been made clear. Readers can judge for themselves whether or not they find the arguments convincing, and whether they can find alternative, perhaps more plausible, sets of assumptions and interpretations under which coherent sets of estimates can be obtained.

A key part of the research programme has been the use of MPES, which allows users to combine information provided in different epidemiological studies which inform different functions of the same underlying set of parameters. This permits us to include more kinds of evidence in each set of estimates, and to check consistency of different kinds of evidence, and to propagate uncertainty in evidence correctly through the model.

Even so, although the use of MPES methods in the report has required some sophisticated modelling, much of the intellectual content of the report has gone into understanding the relationships between the target parameters and what the studies are actually estimating, and to understanding the biases to which the key study designs are vulnerable. The major contribution of the report may, in fact, not be in the estimates, nor even in their coherence, but in the consistent set of interpretations we have proposed for all of the various study designs discussed in the report. These are expressed in the DAGs appearing in each chapter, which make explicit the relationship between what each study estimates and the parameters of interest.

No doubt, future investigators will be able to improve on these characterisations, perhaps based on a better understanding of the same evidence, or perhaps on new evidence. However, we have shown, for the first time, that the clinical and population epidemiology of CT can be effectively studied by developing a formal and explicit analysis of the data-generating mechanisms in the many study designs contributing information.

Summary of findings

1. Two areas of conflict in the work from earlier chapters were identified, and both were resolved by post-hoc re-interpretations of evidence sources and potential biases.

    i. A final estimate of the risk of EP following salpingitis is based on the UK reported rates, multiplied into an estimate of the proportion of EP due to salpingitis, from case–control studies.
    ii. A final estimate of the proportion of TFI attributable to CT, 29%, is derived from our prospective analysis.

2. The public health messages were:

    i. In women aged 16–24 years, annual screening can directly prevent at best 61% of CT-related PID, and 22% of all-cause PID.
    ii. Current guidance for diagnosis and management of PID should not be changed.
    iii. Guidance is required on when women should seek early medical attention to avoid the risk of reproductive damage.
    iv. CT carries a significant risk of reproductive damage, particularly TFI. Every 1000 CT infections in women aged 16–44 years, on average, gives rise to approximately 171 episodes of PID and 73 of salpingitis, 2.0 EPs and 5.1 women with TFI at age 44.

3. Research recommendations:

    i. Serological case–control studies are needed to estimate the proportions of PID, EP and TFI that are caused by CT.
    ii. The same studies could provide information on the relation between age and PEF, on the proportion of PID that is diagnosed, and on the microbiological basis of salpingitis. They would also allow better quantification of the overlap between routine data sources on PID, leading to better estimates of PID incidence.
    iii. Further dynamic modelling should be undertaken, using the estimates of the sequelae of CT infection from this report, to assess whether or not CT screening is cost-effective. This should use a MPES approach to uncertainty propagation and inference, which may require development of new methods of Bayesian computation.

Acknowledgements

Contributions of authors

Dr Malcolm J Price (Research Associate, co-investigator) had a leading role in interpreting literature and study design, performed the data extraction, conceptualised and carried out the statistical analyses, drafted sections of the report, and commented on drafts of the report.

Professor AE Ades (Professor of Public Health Science, Principal Investigator) conceptualised and oversaw the project, had a leading role in interpreting the literature and study design, and drafted the report.

Dr Kate Soldan (Scientist Advisor, co-investigator) contributed to the study design, provision and interpretation of routine data and interpretation of literature, and commented on drafts of the report.

Dr Nicky J Welton (Reader in Evidence Synthesis, co-investigator) contributed to the study design, advised on parts of the statistical analysis, and commented on drafts of the report.

Professor John Macleod (Professor in Clinical Epidemiology and Primary Care, co-investigator) contributed to the study design and commented on drafts of the report.

Dr Ian Simms (Epidemiologist, co-investigator) contributed to the interpretation of literature and commented on drafts of the report.

Dr Daniela DeAngelis (MRC Programme Leader, co-investigator) advised on parts of the statistical analysis and commented on drafts of the report.

Dr Katherine ME Turner (Research Fellow and Senior Lecturer in Veterinary Infectious Diseases) contributed to interpretation of the literature and commented on drafts of the report.

Dr Paddy J Horner (Consultant Senior Lecturer, co-investigator) was the main advisor on clinical and diagnostic issues, made a significant contribution to the study design, drafted sections of the report, had a leading role in interpretation of literature, and commented on drafts of the report.

Contributions of others

Document preparation

We acknowledge the assistance of Katrina Crook and Stephanie Roberts in compiling the material in the report and formatting it for submission.

Scientific input

The authors take full responsibility for study inclusion, interpretation and data extraction, but express their gratitude to Jan van Bergen, Charlotte Gaydos, William Geisler, Pippa Oakeshott and Barbara Van Der Pol, who served as an expert panel to advise on the completeness of literature searches on CT duration in *Chapter 4*.

We thank Sarah Woodhall for assistance in providing and interpreting routine data.

Institutional support

We thank the Medical Research Council (MRC) who funded this research through the Infection and Immunity Board under grant G0801947. MP also received funds from Economic and Social Research Council Postgraduate Fellowship ES/J002909/1 for part of the work.

The authors thank the GPRD for providing data for this study, and all the practices that contribute to the GPRD.

We thank all clinics that report KC-60 data and contributors to the GUMCAD.

We are grateful to the Health and Social Care Information Centre who make HES data available for researchers.

We thank the HPA (now Public Health England) and all the reporters of CT screening data for the provision of data on CT screening outside of GUM clinic.

Publications

Price MJ, Ades AE, Welton NJ, Macleod J, Turner K, Simms I, *et al.* How much tubal factor infertility is caused by chlamydia? Estimation based on serological evidence adjusted for sensitivity and specificity. *Sex Transm Dis* 2012;**39**:608–13.

Price MJ, Ades AE, De Angelis D, Welton NJ, Macleod J, Soldan K, *et al.* Use of latent class models to explain heterogeneity in studies of duration of *Chlamydia trachomatis* infection. *Stat Med* 2013;**32**:1547–60.

Price MJ, Ades AE, De Angelis D, Welton NJ, Macleod J, Soldan K, *et al.* Risk of pelvic inflammatory disease following *Chlamydia trachomatis* infection: analysis of prospective studies with a multistate model. *Am J Epidemiol* 2013;**178**:484–92.

Price MJ, Ades AE, De Angelis D, Welton NJ, Macleod J, Turner K, *et al.* Incidence of *Chlamydia trachomatis* infection in women in England: two methods of estimation. *Epidemiol Infect* 2014;**142**:562–76.

Data sharing statement

The WINBUGS code and data for the analyses in this HTA report are all shown in the appendices. Copies of the WinBUGS code and data as Word documents can be downloaded from www.bristol.ac.uk/social-community-medicine/projects/mpes/ide/chlamydia/.

References

1. Adams EJ, Turner KME, Edmunds WJ. The cost effectiveness of opportunistic chlamydia screening in England. *Sex Transm Infect* 2007;**83**:267–75. http://dx.doi.org/10.1136/sti.2006.024364
2. Roberts TE, Robinson S, Barton PM, Bryan S, McCarthy A, Macleod J, *et al.* Cost effectiveness of home based population screening for *Chlamydia trachomatis* in the UK: economic evaluation of chlamydia screening studies (ClaSS) project. *BMJ* 2007;**335**:291–8. http://dx.doi.org/10.1136/bmj.39262.683345.AE
3. Public Health England (PHE). *National Chlamydia Screening Programme*. URL: www.chlamydiascreening.nhs.uk/ps/index.asp. 2014. (6 October 2015).
4. Stephenson J. Screening for genital chlamydial infection. *Br Med Bull* 1998;**5**:891–902. http://dx.doi.org/10.1093/oxfordjournals.bmb.a011736
5. Low N. Screening programmes for chlamydial infection: when will we ever learn? *BMJ* 2007;**334**:725–8. http://dx.doi.org/10.1136/bmj.39154.378079.BE
6. Wilson J, Junger G. *Principles and Practice of Screening for Disease*. Geneva: WHO; 1968.
7. Chief Medical Officer's Expert Advisory Group. *Main Report of the CMO's Expert Advisory Committee on* Chlamydia trachomatis. London: Department of Health; 1998.
8. Townshend JRP, Turner HS. Analysing the effectiveness of Chlamydia screening. *J Oper Res Soc* 2000;**51**:812–24. http://dx.doi.org/10.1057/palgrave.jors.2600978
9. Department of Health (DH). *The National Strategy for Sexual Health and HIV*. London: DH; 2001.
10. Public Health England (PHE). *The UK National Screening Committee Policy on Chlamydia Screening in Adults*. 2006. URL: http://legacy.screening.nhs.uk/chlamydia-adult (accessed 6 October 2015).
11. Sheringham J, Baraitser P, Simms I. Chlamydia screening in England: a qualitative study of the narrative behind the policy. *BMC Publ Health* 2012;**12**.
12. Ostergaard L, Andersen B, Moller JK, Olesen F. Home sampling versus conventional swab sampling for screening of *Chlamydia trachomatis* in women: a cluster-randomized 1-year follow-up study. *Clin Infect Dis* 2000;**31**:951–7. http://dx.doi.org/10.1086/318139
13. Scholes D, Stergachis A, Heidrich FE, Andrilla H, Holmes KK, Stamm WE. Prevention of pelvic inflammatory disease by screening for cervical chlamydial infection. *N Engl J Med* 1996;**334**:1362–6. http://dx.doi.org/10.1056/NEJM199605233342103
14. Abter EIM. Screening for Chlamydia to prevent pelvic inflammatory disease. *N Engl J Med* 1996;**335**:1531–3. http://dx.doi.org/10.1056/NEJM199611143352012
15. Low N, McCarthy A, Macleod J, Salisbury C, Campbell R, Roberts TE, *et al.* Epidemiological, social, diagnostic, and economic evaluation of population screening for genital chlamydial infection. *Health Technol Assess* 2007;**11**(8). http://dx.doi.org/10.3310/hta11080
16. European Centre for Disease Control and Prevention (ECDC). *Review of Chlamydia Control Activities in EU Countries: Technical Report*. Stockholm: ECDC; 2008.
17. Nelson H, Helfand M. Screening for chlamydial infection. *Am J Prev Med* 2001;**20**:95–107. http://dx.doi.org/10.1016/S0749-3797(01)00253-7
18. Low N, Bender N, Nartey L, Shang A, Stephenson JM. Effectiveness of chlamydia screening: systematic review. *Int J Epidemiol* 2009;**38**:435–48. http://dx.doi.org/10.1093/ije/dyn222

19. Oakeshott P, Kerry S, Aghaizu A, Atherton H, Hay S, Taylor-Robinson D, *et al.* Randomised controlled trial of screening for *Chlamydia trachomatis* to prevent pelvic inflammatory disease: the POPI (prevention of pelvic infection) trial. *BMJ* 2010;**340**:c1642. http://dx.doi.org/10.1136/bmj.c1642

20. Gottlieb S, Berman S, Low N. Screening and treatment to prevent sequelae in women with *Chlamydia trachomatis* genital infection: how much do we know? *J Infect Dis* 2010;**201**:156–67. http://dx.doi.org/10.1086/652396

21. Sheringham J. Screening for chlamydia. *BMJ* 2010;**340**:c1698. http://dx.doi.org/10.1136/bmj.c1698

22. Van den Broek IVF, van Bergen JEAM, Brouwers EEHG, *et al.* Effectiveness of yearly, register based screening for chlamydia in the Netherlands: controlled trial with randomised stepped wedge implementation. *BMJ* 2012;344. http://dx.doi.org/10.1136/bmj.e4316

23. Anderson B, van Valkengoed I, Skolowski I, Moller J, Ostergaard L, Olesen F. Impact of intensified testing for urogenital *Chlamydia trachomatis* infections: a randomised study with 9-year follow-up. *Sex Transm Infect* 2010;**87**:156–61. http://dx.doi.org/10.1136/sti.2010.042192

24. Soldan K, Berman S. Danish health register study: a randomised trial with findings about the implementation of chlamydia screening, but not about its benefits. *Sex Transm Infect* 2011;**87**:86–7. http://dx.doi.org/10.1136/sti.2010.048082

25. Hocking J. Protocol 12PRT/9010:Australian Chlamydia Control Effectiveness Pilot (ACCEPt): a cluster randomised controlled trial of chlamydia testing in general practice (ACTRN1260000297022). URL: www.thelancet.com/protocol-reviews/12PRT-9010. 2010 (accessed 6 October 2015).

26. Kretzschmar M, Turner KME, Barton PM, Edmunds WJ, Low N. Predicting the population impact of chlamydia screening programmes: comparative mathematical modelling study. *Sex Transm Infect* 2009;**85**:359–66. http://dx.doi.org/10.1136/sti.2009.036251

27. Althaus CL, Turner KME, Schmid BV, Heijne JCM, Kretzschmar M, Low N. Transmission of *Chlamydia trachomatis* through sexual partnerships: a comparison between three individual-based models and empirical data. *J R Soc Interface* 2012;**9**:136–46. http://dx.doi.org/10.1098/rsif.2011.0131

28. Brunham RC, Rey-Ladino J. Immunology of Chlamydia infection: implications for a *Chlamydia trachomatis* vaccine. *Nat Rev Immunol* 2005;**5**:149–61. http://dx.doi.org/10.1038/nri1551

29. Hu D, Hook EW, Goldie SJ. Screening for *Chlamydia trachomatis* in women 15 to 29 years of age: a cost-effectiveness analysis. *Ann Int Med* 2004;**141**:501–13. http://dx.doi.org/10.7326/0003-4819-141-7-200410050-00006

30. Hu D, Hook EW, Goldie SJ. The impact of natural history parameters on the cost-effectiveness of *Chlamydia trachomatis* screening strategies. *Sex Transm Dis* 2006;**33**:428–36. http://dx.doi.org/10.1097/01.olq.0000200577.46261.b0

31. Paavonen J, Eggert-Kruse W. *Chlamydia trachomatis*: impact on human reproduction. *Hum Reprod Update* 1999;**5**:433–47. http://dx.doi.org/10.1093/humupd/5.5.433

32. Welte R, Postma M, Leidl R, Kretzschmar M. Costs and effects of chlamydial screening: dynamic versus static modeling. *Sex Transm Dis* 2005;**32**:474–83. http://dx.doi.org/10.1097/01.olq.0000161181.48687.cf

33. Westrom L, Joesoef R, Reynolds G, Hagdu A, Thompson SE. Pelvic inflammatory disease and fertility: a cohort study of 1,844 women with laparoscopically verified disease and 657 control women with normal laparoscopic results. *Sex Transm Dis* 1992;**19**:185–92. http://dx.doi.org/10.1097/00007435-199207000-00001

34. Westrom LV. Sexually-transmitted diseases and infertility. *Sex Transm Dis* 1994;**21**:S32–7.

35. Cahill DJ, Wardle PG. Management of infertility. *BMJ* 2002;**325**:28–32. http://dx.doi.org/10.1136/bmj.325.7354.28

36. Low N, Egger M, Sterne JAC, Harbord RM, Ibrahim F, Lindblom B, *et al.* Incidence of severe reproductive tract complications associated with diagnosed genital chlamydial infection: the Uppsala Women's Cohort Study. *Sex Transm Infect* 2006;**82**:212–18. http://dx.doi.org/10.1136/sti.2005.017186

37. Simms I, Stephenson JM, Mallinson H, Peeling RW, Thomas K, Gokhale R, *et al.* Risk factors associated with pelvic inflammatory disease. *Sex Transm Infect* 2006;**82**:452–7. http://dx.doi.org/10.1136/sti.2005.019539

38. Gottlieb SL, Martin DH, Xu FJ, Byrne GI, Brunham RC. Summary: the natural history and immunobiology of *Chlamydia trachomatis* genital infection and implications for Chlamydia control. *J Infect Dis* 2010;**201**:S190–204. http://dx.doi.org/10.1086/652401

39. Van Valkengoed IGM, Morre SA, van den Brule AJC, Meijer CJLM, Bouter LM, Boeke AJP. Overestimation of complication rates in evaluations of *Chlamydia trachomatis* screening programmes – implications for cost-effectiveness analyses. *Int J Epidemiol* 2004;**33**:416–25. http://dx.doi.org/10.1093/ije/dyh029

40. Simms I, Horner P. Has the incidence of pelvic inflammatory disease following chlamydial infection been overestimated? *Int J STD AIDS* 2008;**19**:285–6. http://dx.doi.org/10.1258/ijsa.2007.007316

41. Guyatt G, Sackett D, Sinclair J, Hayward R, Cook D, Cook R. Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. *JAMA* 1885;**274**:1800–04. http://dx.doi.org/10.1001/jama.1995.03530220066035

42. Ades AE, Sutton AJ. Multiparameter evidence synthesis in epidemiology and medical decision making: current approaches. *Stat Soc* 2006;**169**:5–35. http://dx.doi.org/10.1111/j.1467-985X.2005.00377.x

43. Ades AE, Welton N, Caldwell D, Price M, Goubar A, Lu G. Multiparameter evidence synthesis in epidemiology and medical decision-making. *J Health Serv Res Policy* 2008;**13**:12–22. http://dx.doi.org/10.1258/jhsrp.2008.008020

44. Spiegelhalter D, Thomas A, Best N, Lunn D. *WinBUGS User Manual Version 1.4, January 2003*. Upgraded to Version 1.4.3, 2007. (www.mrc-bsu.cam.ac.uk/bugs)

45. Birge R. Probable values of e, h, e/m, and a. *Phys Rev* 1932;**40**:228–61. http://dx.doi.org/10.1103/PhysRev.40.228

46. Brand KP, Small MJ. Updating uncertainty in an integrated risk assessment: conceptual framework and methods. *Risk Anal* 1995;**15**:719–31. http://dx.doi.org/10.1111/j.1539-6924.1995.tb01344.x

47. Poole D, Raftery AE. Inference for deterministic simulation models: the Bayesian melding approach. *J Am Stat Assoc* 2000;**95**:1244–55. http://dx.doi.org/10.1080/01621459.2000.10474324

48. Raftery AE, Givens GH, Zeh JE. Inference from a deterministic population dynamics model for Bowhead whales (with discussion). *J Am Stat Assoc* 1995;**90**:402–30. http://dx.doi.org/10.1080/01621459.1995.10476529

49. Hasselblad V, McCrory DC. Meta-analytic tools for medical decision making: a practical guide. *Med Decis Making* 1995;**15**:81–96. http://dx.doi.org/10.1177/0272989X9501500112

50. Eddy DM, Hasselblad V, Shachter R. *Meta-analysis by the Confidence Profile Method*. London: Academic Press; 1992.

51. Sutton AJ, Abrams KR, Ades AE, Cooper NJ, Welton NJ. *Evidence Synthesis for Decision Making*. Chichester: Wiley; 2008.

52. Goubar A, Ades AE, De Angelis D, Tom BD, Birrell PJ, Charlett A, *et al.* Estimates of human immunodeficiency virus prevalence and proportion diagnosed based on Bayesian multiparameter synthesis of surveillance data. *J Roy Stat Soc A* 2008;**171**:541–80. http://dx.doi.org/10.1111/j.1467-985X.2007.00537.x

53. Presanis A, Pebody RG, Paterson BJ, Tom BD, Birrell PJ, Charlett A, *et al.* Changes in the severity of the 2009 pandemic A/H1N1 influenza in England: a Bayesian evidence synthesis. *BMJ* 2011;**343**:d5408. http://dx.doi.org/10.1136/bmj.d5408

54. Presanis AM, Gill ON, Chadborn TR, Hill C, Hope V, Logan L, *et al.* Insights into the rise in HIV infections in England and Wales, 2001 to 2008: a Bayesian synthesis of prevalence studies. *Aids* 2010;**24**:2849–58. http://dx.doi.org/10.1097/QAD.0b013e32834021ed

55. Sweeting MJ, De Angelis D, Hickman D, Ades AE. Estimating HCV prevalence in England and Wales by synthesising evidence from multiple data sources: assessing data conflict and model fit. *Biostatistics* 2008;**9**:715–34. http://dx.doi.org/10.1093/biostatistics/kxn004

56. Colbourn TE, Asseburg C, Bojke L, Phillips Z, Welton NJ, Claxton A, *et al.* Preventive strategies for group B streptococcal and other bacterial infections in early infancy: cost effectiveness and value of information analyses. *BMJ* 2007;**335**:655–61. http://dx.doi.org/10.1136/bmj.39325.681806.AD

57. Ades AE, Cliffe S. Markov Chain Monte Carlo estimation of a multi-parameter decision model: consistency of evidence and the accurate assessment of uncertainty. *Medical Decision Making* 2002;**22**:359–71. http://dx.doi.org/10.1177/027298902400448920

58. Cooper H, Hedges L. *The Handbook of Research Synthesis*. New York, NY: Russell Sage Foundation; 1994.

59. Sutton AJ, Higgins JPT. Recent developments in meta-analysis. *Stat Med* 2008;**27**:625–50. http://dx.doi.org/10.1002/sim.2934

60. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.0* [updated February 2008]. Chichester: The Cochrane Collaboration, Wiley; 2008.

61. Land JA, van Bergen JEAM, Morre SA, Postma MJ. Epidemiology of *Chlamydia trachomatis* infection in women and the cost-effectiveness of screening. *Hum Reprod Update* 2010;**16**:189–204. http://dx.doi.org/10.1093/humupd/dmp035

62. Simms I, Stephenson JM. Pelvic inflammatory disease epidemiology: what do we know and what do we need to know? *Sex Transm Infect* 2000;**76**:80–7. http://dx.doi.org/10.1136/sti.76.2.80

63. Centers for Disease Control and Prevention (CDC). Sexually Transmitted Diseases Treatment Guidelines, 2010. *MMWR* 2010;**59**:1–116.

64. Horner P, Boag F. *UK National Guideline for the Management of Genital Tract Infection with* Chlamydia trachomatis. 2006. URL: http://www.bashh.org/documents/65.pdf (accessed 6 October 2015).

65. Black CM. Current methods of laboratory diagnosis of *Chlamydia trachomatis* infections. *Clin Microbiol Rev* 1997;**10**:160–84.

66. Cameron ST, Glasier A, Scott G, Young H, Melvin L, Johnstone A, *et al.* Novel interventions to reduce re-infection in women with chlamydia: a randomized controlled trial. *Hum Reprod* 2009;**24**:888–95. http://dx.doi.org/10.1093/humrep/den475

67. Cameron ST, Glasier A, Muir A, Scott G, Johnstone A, Quarrell H, *et al.* Expedited partner therapy for *Chlamydia trachomatis* at the community pharmacy. *BJOG* 2010;**117**:1074–9. http://dx.doi.org/10.1111/j.1471-0528.2010.02573.x

68. Dombrowski JC, Golden MR. Accelerated partner therapy: a promising new partner treatment option. *Sex Transm Infect* 2012;**88**:2–3. http://dx.doi.org/10.1136/sextrans-2011-050377

69. Estcourt C, Sutcliffe L, Cassell J, Mercer CH, Copas A, James L, *et al.* Can we improve partner notification rates through expedited partner therapy in the UK? Findings from an exploratory trial of Accelerated Partner Therapy (APT). *Sex Transm Infect* 2012;**88**:21–6. http://dx.doi.org/10.1136/sti.2010.047258

70. Golden MR, Whittington WL, Handsfield HH, Hughes JP, Stamm WE, Hogben M, *et al.* Effect of expedited treatment of sex partners on recurrent or persistent gonorrhea or chlamydial infection. *N Engl J Med* 2005;**352**:676–85. http://dx.doi.org/10.1056/NEJMoa041681

71. Shiely FPHD, Hayes KPHD, Thomas KM, Kerani RP, Hughes JP, Whittington WL, *et al.* Expedited partner therapy: a robust intervention. *Sex Transm Dis* 2010;**37**:602–7. http://dx.doi.org/10.1097/olq.0b013e3181e1a296

72. Van den Broek IV, Hoebe CJ, van Bergen JE, Brouwers EE, de Feijter EM, Fennema JS, *et al.* Evaluation design of a systematic, selective, internet-based, Chlamydia screening implementation in the Netherlands, 2008–2010: implications of first results for the analysis. *BMC Infect Dis* 2010;**10**:89. http://dx.doi.org/10.1186/1471-2334-10-89

73. Geisler WM, Wang C, Morrison SG, Black CM, Bandea CI, Hook EW. The natural history of untreated *Chlamydia trachomatis* infection in the interval between screening and returning for treatment. *Sex Transm Dis* 2008;**35**:119–23. http://dx.doi.org/10.1097/OLQ.0b013e318151497d

74. Geisler WM. Duration of untreated, uncomplicated *Chlamydia trachomatis* genital infection and factors associated with chlamydia resolution: a review of human studies. *J Infect Dis* 2010;**201**:S104–13. http://dx.doi.org/10.1086/652402

75. Cates W, Steiner M. Dual protection against unintended pregnancy and sexually transmitted infections: what is the best contraceptive approach? *Sex Transm Dis* 2002;**293**:168–74. http://dx.doi.org/10.1097/00007435-200203000-00007

76. Althaus C, Heijne J, Low N. Towards more robust estimates of the transmissibility of *Chlamydia trachomatis*. *Sex Transm Dis* 2012;**39**:402–4. http://dx.doi.org/10.1097/OLQ.0b013e318248a550

77. Faculty of Sexual and Reproductive Healthcare. *Emergency Contraception*. London: Royal College of Obstetricians and Gynaecologists; 2012.

78. Faculty of Forensic and Legal Medicine. *Recommendations for the Collection of Forensic Specimens from Complainants and Suspects.* London: Royal College of Physicians; 2014.

79. Joyner JL, Douglas JM, Foster M, Judson FN. Persistence of *Chlamydia trachomatis* infection detected by polymerase chain reaction in untreated patients. *Sex Transm Dis* 2002;**29**:196–200. http://dx.doi.org/10.1097/00007435-200204000-00002

80. Geisler W, Lensing S, Press C. Spontaneous resolution of genital *Chlamydia trachomatis* Infection in women and protection from reinfection. *J Infect Dis* 2013;**207**:1850–56. http://dx.doi.org/10.1093/infdis/jit094

81. Macleod J, Salisbury C, Low N, McCarthy A, Sterne JA, Holloway A, *et al.* Coverage and uptake of systematic postal screening for genital *Chlamydia trachomatis* and prevalence of infection in the United Kingdom general population: cross sectional study. *BMJ* 2005;**330**:940–2. http://dx.doi.org/10.1136/bmj.38413.663137.8F

82. Paavonen J, Westrom L, Eschenbach DA. Pelvic Inflammatory Disease. InHolmes K, Sparling PF, Stamm WE, editors. *Sexually Transmitted Disease*. 4th edn. London: McGraw Hill; 2008:1021–2.

83. British Association for Sexual Health and HIV (BASHH) Clinical Effectiveness Group. *UK National Guideline for the Management of Pelvic Inflammatory Disease 2011*. Macclesfield: BASHH; 2011.

84. Hillis SD, Joesoef R, Marchbanks PA, Wasserheit JN, Cates W, Westrom L. Delayed care of pelvic inflammatory disease as a risk factor for impaired fertility. *Am J Obstet Gynecol* 1993;**168**:1503–9. http://dx.doi.org/10.1016/S0002-9378(11)90790-X

85. Treharne JD, Ripa KT, Mardh PA, Svensson L, Westrom L, Darougar S. Antibodies to *Chlamydia trachomatis* in acute salpingitis. *Brit J Vener Dis* 1979;**55**:26–9. http://dx.doi.org/10.1136/sti.55.1.26

86. Ness R, Keder L, Soper D. Oral contraception and the recognition of endometritis. *Am J Obstet Gynecol* 1997;**176**:580–5. http://dx.doi.org/10.1016/S0002-9378(97)70551-9

87. Ness RB, Soper DE, Peipert J, Sondheimer SJ, Holley RL, Sweet RL, *et al.* Design of the PID Evaluation and Clinical Health (PEACH) Study. *Control Clin Trials* 1998;**19**:499–514. http://dx.doi.org/10.1016/S0197-2456(98)00022-1

88. Taylor-Robinson D, Stacey CM, Jensen JS, Thomas BJ, Munday PE. Further observations, mainly serological, on a cohort of women with or without pelvic inflammatory disease. *Int J STD AIDS* 2009;**20**:712–18. http://dx.doi.org/10.1258/ijsa.2008.008489

89. Hager W, Eschenbach D, Spence M. Criteria for diagnosis and grading of salpingitis. *Obstet Gynecol* 1983;**61**:113–14.

90. French CE, Hughes G, Nicholson A, Jung M, Ross JD, Williams T, *et al.* Estimation of the rate of pelvic inflammatory disease diagnoses: trends in England, 2000–2008. *Sex Transm Dis* 2011;**38**:158–62. http://dx.doi.org/10.1097/OLQ.0b013e3181f22f3e

91. Horner P, Blee K, Adams E. Time to manage *Mycoplasma genitalium* as an STI: but not with azithromycin 1 gram! *Current Opinion in Infectious Diseases* 2014;**27**:68–74. http://dx.doi.org/10.1097/QCO.0000000000000030

92. Brotman R, Ravel J, Bavoil P. Microbiome, sex hormones, and immune responses in the reproductive tract: challenges for vaccine development against sexually transmitted infections. *Vaccine* 2014;**32**:1543–52. http://dx.doi.org/10.1016/j.vaccine.2013.10.010

93. Mehta S. Systematic review of randomized trials of treatment of male sexual partners for improved bacterial vaginosis outcomes in women. *Sex Transm Dis* 2012;**39**:822–30. http://dx.doi.org/10.1097/OLQ.0b013e3182631d89

94. Koumans EH, Sternberg M, Bruce C, McQuillan G, Kenddrick J, Sutton M, *et al.* The prevalence of bacterial vaginosis in the United States, 2001–2004; associations with symptoms, sexual behaviors, and reproductive health. *Sex Transm Dis* 2007;**34**:864–9. http://dx.doi.org/10.1097/OLQ.0b013e318074e565

95. British Association for Sexual Health and HIV (BASHH) Clinical Effectiveness Group. *UK National Guideline for the Management of Bacterial Vaginosis*. Macclesfield: BASHH; 2012.

96. Martin DH. The microbiota of the vagina and its influence on women's health and disease. *Am J Med Sci* 2012;**343**:2–9. http://dx.doi.org/10.1097/MAJ.0b013e31823ea228

97. Nicholson A, Rait G, Murray-Thomas T, Hughes G, Mercer C, Cassell J. Management of first episode pelvic inflammatory disease in primary care: results form a large primary care database. *Br J Gen Pract* 2010;**60**. http://dx.doi.org/10.3399/bjgp10x532404

98. Haggerty C, Gottlieb S, Taylor B, Low N, Xu F, Ness R. Risk of sequelae after *Chlamydia trachomatis* genital infection in women. *J Infect Dis* 2010;**201**:S134–55. http://dx.doi.org/10.1086/652395

99. Sellors JW, Mahony JB, Chernesky MA, Rath DJ. Tubal factor infertility: an association with prior chlamydial Infection and asymptomatic salpingitis. *Fertil Steril* 1988;**49**:451–7.

100. Cates W, Joesoef MR, Goldman MB. Atypical pelvic inflammatory disease: can we identify clinical predictors. *Am J Obstet Gynecol* 1993;**169**:341–6. http://dx.doi.org/10.1016/0002-9378(93)90085-W

101. Svensson L, Westrom L, Ripa KT, Mardh PA. Differences in some clinical and laboratory parameters in acute salpingitis related to culture and serologic findings. *Am J Obstet Gynecol* 1980;**138**:1017–21.

102. Osser S, Persson K, Liedholm P. Tubal infertility and silent chlamydial salpingitis. *Hum Reprod* 1989;**4**:280–4.

103. Wolner-Hanssen P. Silent pelvic inflammatory disease: is it overstated? *Obstet Gynecol* 1995;**86**:321–5. http://dx.doi.org/10.1016/0029-7844(95)00177-S

104. National Institute for Health and Care Excellence (NICE). *Pelvic Inflammatory Disease. Clinical Knowledge Summary*. 2013. URL: http://cks.nice.org.uk/pelvic-inflammatory-disease#!topicsummary (accessed 6 October 2015).

105. Lazaro N. *Sexually Transmitted Infections in Primary Care 2013*. RCGP/BASHH. 2nd edn. London: Royal College of General Practitioners; 2013. (www.rcgp.org and www.bashh.org/guidelines)

106. Adams EJ, Charlett A, Edmunds WJ, Hughes G. *Chlamydia trachomatis* in the United Kingdom: a systematic review and analysis of prevalence studies. *Sex Transm Infect* 2004;**80**:354–62. http://dx.doi.org/10.1136/sti.2003.005454

107. Hillis SD, Owens LM, Marchbanks PA, Amsterdam LF, MacKenzie WR. Recurrent chlamydial infections increase the risks of hospitalization for ectopic pregnancy and pelvic inflammatory disease. *Am J Obstet Gynecol* 1997;**176**:103–7. http://dx.doi.org/10.1016/S0002-9378(97)80020-8

108. Eschenbach DA, Buchanan TM, Pollock HM, Forsyth PS, Alexander ER, Linn JS, *et al.* Polymicrobial etiology of acute pelvic inflammatory disease. *N Engl J Med* 1975;**293**:166–71. http://dx.doi.org/10.1056/NEJM197507242930403

109. Allsworth JE, Peipert JF. Prevalence of bacterial vaginosis: 2001–2004 National Health and Nutrition Examination Survey data. *Obstet Gynecol* 2007;**109**:114–20. http://dx.doi.org/10.1097/01.AOG.0000247627.84791.91

110. Westrom L, Bengtsson LP, Mardh PA. Incidence, trends, and risks of ectopic pregnancy in a population of women. *Br Med J* 1981;**282**:15–18. http://dx.doi.org/10.1136/bmj.282.6257.15

111. Peipert J, Ness R, Blume J. Clinical predictors of endometritis in women with symptoms and signs of pelvic inflammatory disease. *Am J Obstet Gynecol* 2001;**184**:856–63. http://dx.doi.org/10.1067/mob.2001.113847

112. Sellors J, Mahony J, Goldsmith C, Rath D, Mander R, Hunter B, *et al.* The accuracy of clinical findings and laparoscopy in pelvic inflammatory disease. *Am J Obstet Gynecol* 1991;**164**:113–20. http://dx.doi.org/10.1016/0002-9378(91)90639-9

113. Westrom L, Wolner-Hanssen P. Pathogenesis of pelvic inflammatory disease. *Genitourin Med* 1993;**69**:9–17. http://dx.doi.org/10.1136/sti.69.1.9

114. Short VL, Totten PA, Ness RB, Astete SG, Kelsey SF, Haggerty CL. Clinical presentation of *Mycoplasma genitalium* infection versus *Neisseria gonorrhoeae* infection among women with pelvic inflammatory disease. *Clin Infect Dis* 2009;**48**:41–7. http://dx.doi.org/10.1086/594123

115. Taylor B, Ness R, Darville T. Microbial correlates of delayed care for pelvic inflammatory disease. *Sex Transm Dis* 2011;**387**:434–8. http://dx.doi.org/10.1097/OLQ.0b013e3181ffa7c7

116. Farquhar CM. Ectopic pregnancy. *Lancet* 2005;**366**:583–91. http://dx.doi.org/10.1016/S0140-6736(05)67103-6

117. Health & Social Care Information Centre (HSCIC). *Hospital Episode Statistics* 2012. URL: www.hscic.gov.uk/hes (accessed 11 December 2012).

118. Office for National Statistics (ONS). *Conception Statistics, England and Wales*. 2012. URL: www.ons.gov.uk/ons/publications/all-releases.html?definition = tcm%3A77–27423 (accessed 10 October 2012).

119. Coste J, Jobspira N, Aubletcuvelier B, Germain E, Glowackzower E, Fernandez H, *et al.* Incidence of ectopic pregnancy: first results of a population-based register in France. *Hum Reprod* 1994;**9**:742–5.

120. Hoover KW, Tao G, Kent CK. Trends in the diagnosis and treatment of ectopic pregnancy in the United States. *Obstet Gynecol* 2010;**115**:495–502. http://dx.doi.org/10.1097/AOG.0b013e3181d0c328

121. Egger M, Low N, Smith GD, Lindblom B, Herrmann B. Screening for chlamydial infections and the risk of ectopic pregnancy in a county in Sweden: ecological analysis. *BMJ* 1998;**316**(7147):1776–80. http://dx.doi.org/10.1136/bmj.316.7147.1776

122. Trabert B, Holt VL, Yu O, Van den Eeden SK, Scholes D. Population-based ectopic pregnancy trends, 1993–2007. *Am J Prev MEd* 2011;**40**:556–60. http://dx.doi.org/10.1016/j.amepre.2010.12.026

123. Simms I, Rogers PA, Nicoll A. The influence of demographic change and cumulative risk of pelvic inflammatory disease on the incidence of ectopic pregnancy. *Epidemiol Infect* 1997;**119**:49–52. http://dx.doi.org/10.1017/S0950268897007711

124. Westrom L. Effect of pelvic inflammatory disease on fertility. *Venereology* 1995;**8**:219–22.

125. National Collaborating Centre for Women's and Children's Health. *Fertility: Assessment and Treatment for People with Fertility Problems*. London: Royal College of Obstetricians and Gynaecologists; 2013.

126. Templeton A, Fraser C, Thompson B. The Epidemiology of infertility in Aberdeen. *BMJ* 1990;**301**:148–52. http://dx.doi.org/10.1136/bmj.301.6744.148

127. Bhattacharya S, Porter M, Amalraj E, Templeton A, Hamilton M, Lee AJ, *et al.* The epidemiology of infertility in the North East of Scotland. *Hum Reprod* 2009;**24**:3096–107. http://dx.doi.org/10.1093/humrep/dep287

128. Gunnell DJ, Ewings P. Infertility prevalence, needs assessment and purchasing. *J Publ Health Med* 1994;**16**:29–35.

129. Oakley L, Doyle P, Maconochie N. Lifetime prevalence of infertility and infertility treatment in the UK: results from a population-based survey of reproduction. *Hum Reprod* 2008;**23**:447–50. http://dx.doi.org/10.1093/humrep/dem369

130. Human Fertilisation and Embryology Authority (HFEA). *Facts and Figures. Long-term Trends Data – Patients Treated.* 2011. URL: www.hfea.gov.uk/2585.html (accessed 7 October 2015).

131. Human Fertilisation and Embryology Authority. *Facts and Figures 2006: Fertility Problems and Treatment*. London: HFEA; 2008.

132. Howard FM. Chronic Pelvic Pain. *Obstet Gynecol* 2003;**101**:594–611. http://dx.doi.org/10.1016/S0029-7844(02)02723-0

133. Thorne C. *UK National Screening Committee Policy Review: Chlamydia Screening in Pregnancy: an Evidence Review*. London: National Screening Committee; 2011. (www.screening.nhs.uk/chlamydia-pregnancy)

134. Kaltenthaler E, Tappenden P, Paisley S, Squires H. *Identifying and Reviewing Evidence to Inform the Conceptualisation and Population of Cost-effectiveness Models*. NICE Decision Support Unit, Technical Support Document 13. Sheffield: University of Sheffield; 2011.

135. Pirolli P, Card S. Information foraging. *Psychol Rev* 1999;**106**:643–75. http://dx.doi.org/10.1037/0033-295X.106.4.643

136. Glanville J, Paisley S. Searching for Evidence for Cost-Effectiveness Decisions. In Shemilt I, Mugford M, Vale L, Marsh K, Donaldson C, editors. *Evidence-based Decisions and Economics: Health Care, Social Welfare, Education and Criminal Justice*. Oxford: Wiley-Blackwell; 2010. http://dx.doi.org/10.1002/9781444320398.ch7

137. Kretzschmar M, Welte R, Van den Hoek A, Postma MJ. Comparative model-based analysis of screening programs for *Chlamydia trachomatis* infections. *Am J Epidemiol* 2001;**153**:90–101. http://dx.doi.org/10.1093/aje/153.1.90

138. Kretzschmar M, van Duynhoven YT, Severijnen AJ. Modeling prevention strategies for gonorrhea and Chlamydia using stochastic network simulations. *Am J Epidemiol* 1996;**144**:306–17. http://dx.doi.org/10.1093/oxfordjournals.aje.a008926

139. Golden MR, Schillinger JA, Markowitz L, St Louis ME. Duration of untreated genital infections with *Chlamydia trachomatis*: a review of the literature. *Sex Transm Dis* 2000;**27**:329–37. http://dx.doi.org/10.1097/00007435-200007000-00006

140. Turner KME, Adams EJ, Gay N, Ghani AC, Mercer C, Edmunds WJ. Developing a realistic sexual network model of chlamydia transmission in Britain. *Theor Biol Med Model* 2006;**3**:3. http://dx.doi.org/10.1186/1742-4682-3-3

141. Turner KME, Adams EJ, LaMontagne DS, Emmett L, Baster K, Edmunds WJ. Modelling the effectiveness of chlamydia screening in England. *Sex Transm Infect* 2006;**82**:496–502. http://dx.doi.org/10.1136/sti.2005.019067

142. Korenromp EL, Sudaryo MK, de Vlas SJ, Gray RH, Sewankambo NK, Serwadda D, *et al.* What proportion of episodes of gonorrhoea and chlamydia becomes symptomatic? *Int J STD AIDS* 2002;**13**:91–101. http://dx.doi.org/10.1258/0956462021924712

143. Mol BWJ, Dijkman B, Wertheim P, Lijmer J, VanderVeen F, Bossuyt PMM. The accuracy of serum chlamydial antibodies in the diagnosis of tubal pathology: a meta-analysis. *Fertil Steril* 1997;**67**:1031–7. http://dx.doi.org/10.1016/S0015-0282(97)81435-5

144. Risser WL, Risser JMH. The incidence of pelvic inflammatory disease in untreated women infected with *Chlamydia trachomatis*: a structured review. *Int J STD AIDS* 2007;**18**:727–31. http://dx.doi.org/10.1258/095646207782212351

145. Wallace LA, Scouler A, Hart G. What is the excess risk of infertility in women after genital chlamydia infection? A systematic review of the evidence. *Sex Transm Infect* 2008;**84**:171–5. http://dx.doi.org/10.1136/sti.2007.026047

146. Watson EJ, Templeton A, Russell I, Paavonen J, Mardh PA, Stary A, *et al.* The accuracy and efficacy of screening tests for *Chlamydia trachomatis*: a systematic review. *J Med Microbiol* 2002;**51**:1021–31. http://dx.doi.org/10.1099/0022-1317-51-12-1021

147. Ankum WM, Mol BWJ, VanderVeen F, Bossuyt PMM. Risk factors for ectopic pregnancy: a meta-analysis. *Fertil Steril* 1996;**65**:1093–9.

148. Rockhill B, Newmam B, Weinberg C. Use and misuse of population attributable fractions. *Am J Publ Health* 1998;**88**:15–19. http://dx.doi.org/10.2105/AJPH.88.1.15

149. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3rd edn. Philadelphia: Lippincott, Williams & Wilkins; 2008.

150. Spiegelhalter DJ, Abrams KR, Myles J. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. New York: Wiley; 2004.

151. Hastings WK. Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika* 1970;**57**:97–109. http://dx.doi.org/10.1093/biomet/57.1.97

152. Gilks WR, Richardson S, Spiegelhalter DJ. *Markov Chain Monte Carlo in Practice*. London: Chapman & Hall/CRC; 1996.

153. Gelman A, Carlin JG, Stern HS, Rubin DB. *Bayesian Data Analysis*. 1st edn. London: Chapman & Hall; 1995.

154. Geman S, Geman D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1984;**6**:721–41. http://dx.doi.org/10.1109/TPAMI.1984.4767596

155. Brooks SP, Gelman A. Alternative methods for monitoring convergence of iterative simulations. *J Comput Graph Stat* 1998;**7**:434–55.

156. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *J Roy Stat Soc B* 2002;**64**:583–616. http://dx.doi.org/10.1111/1467-9868.00353

157. Dempster AP. The direct use of likelihood for significance testing. *Stat Comput* 1997;**7**:247–52. http://dx.doi.org/10.1023/A:1018598421607

158. Akaike H. Likelihood of a model and information criteria. *J Econometrics* 1981;**16**:3–14. http://dx.doi.org/10.1016/0304-4076(81)90071-3

159. Spiegelhalter DJ, Best N, Carlin BP. *Bayesian Deviance: the Effective Number of Parameters, and the Comparison of Arbitrarily Complex* Models. Technical Report 98–009. Division of Biostatistics. Minneapolis, MN: University of Minnesota; 1998.

160. Lunn D. *WinBUGS Development Interface (WBDev)*. 2004. URL: http://winbugs-development.mrc-bsu.cam.ac.uk/ (accessed 21 March 2013).

161. Lunn D. *WinBUGS Differential Interface (WBDiff)*. 2004. URL: http://winbugs-development.mrc-bsu.cam.ac.uk/ (accessed 21 March 2013).

162. O'Hagan A. HSSS Model Criticism. In Green PJ, editor. *Highly Structured Stochastic Systems*. Oxford: Oxford University Press; 2002.

163. Ohlssen D, Spiegelhalter DJ. *Conflict Diagnosis in Directed Acyclic Graphs*. MRC Public Health Service Research Network Workshop, 2006.

164. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking Consistency in Mixed Treatment Comparison Meta-analysis. *Stat Med* 2010;**29**:932–44. http://dx.doi.org/10.1002/sim.3767

165. Rahm VA, Belsheim J, Gleerup A, Gnarpe H, Rosen G. Asymptomatic carriage of *Chlamydia trachomatis*: a study of 109 teenage girls. *Eur J Sex Transm Dis* 1986;**3**:91–4.

166. Althaus C, Heijne J, Roellin A, Low N. Transmission dynamics of *Chlamydia trachomatis* affect the impact of screening programs. *Epidemics* 2010;**2**:123–31. http://dx.doi.org/10.1016/j.epidem.2010.04.002

167. Jewell NP. Mixtures of exponential distributions. *Ann Stat* 1982;**10**:479–84. http://dx.doi.org/10.1214/aos/1176345789

168. Price MJ, Ades AE, Angelis DD, Welton NJ, Mcleod J, Soldan K, *et al.* Mixture-of-exponentials models to explain heterogeneity in studies of the duration of *Chlamydia trachomatis* infection. *Stat Med* 2013;**32**:1547–60. http://dx.doi.org/10.1002/sim.5603

169. Molano M, Meijer CJLM, Weiderpass E, Arslan A, Posso H, Franceschi S, *et al.* The natural course of *Chlamydia trachomatis* infection in asymptomatic Colombian women: a 5-year follow-up study. *J Infect Dis* 2005;**191**:907–16. http://dx.doi.org/10.1086/428287

170. Ginocchio RHS, Veenstra DL, Connell FA, Marrazzo JM. The clinical and economic consequences of screening young men for genital chlamydial infection. *Sex Transm Dis* 2003;**30**:99–106. http://dx.doi.org/10.1097/00007435-200302000-00001

171. Goeree R, Jang D, Blackhouse G, Chong S, Mahoney J, Sellors J, *et al.* Cost-effectiveness of screening swab or urine specimens for *Chlamydia trachomatis* from young Canadian women in Ontario. *Sex Transm Dis* 2001;**28**:701–9. http://dx.doi.org/10.1097/00007435-200112000-00007

172. Mehta SD, Bishai D, Howell R, Rothman RE, Quinn TC, Zenilman JM. Cost-effectiveness of five strategies for gonorrhea and chlamydia control among female and male emergency department patients. *Sex Transm Dis* 2002;**29**:83–91. http://dx.doi.org/10.1097/00007435-200202000-00004

173. Wang LY, Burstein GR, Cohen DA. An economic evaluation of a school-based sexually transmitted disease screening program. *Sex Transm Dis* 2002;**29**:737–45. http://dx.doi.org/10.1097/00007435-200212000-00001

174. Welte R, Kretzschmar M, Leidl R, Van den Hoek A, Jager JC, Postma MJ. Cost-effectiveness of screening programs for *Chlamydia trachomatis*: a population-based dynamic approach. *Sex Transm Dis* 2000;**27**:518–29. http://dx.doi.org/10.1097/00007435-200010000-00005

175. Rogers SM, Miller WC, Turner CF, Elen J, Zenilman J, Rothman R, *et al.* Concordance of *Chlamydia trachomatis* infections within sexual partnerships. *Sex Transm Infect* 2008;**84**:23–8. http://dx.doi.org/10.1136/sti.2007.027029

176. Parks KS, Dixon PB, Richey CM, Hook EW. Spontaneous clearance of *Chlamydia trachomatis* infection in untreated patients. *Sex Transm Dis* 1997;**24**:229–35. http://dx.doi.org/10.1097/00007435-199704000-00008

177. Schachter J, Hanna L, Hill EC, Massad S, Sheppard CW, Conte JE, Jr, *et al.* Are chlamydial infections the most prevalent venereal disease. *JAMA* 1975;**231**:1252–5. http://dx.doi.org/10.1001/jama.1975.03240240022018

178. Tsukagoshi T. Epidemiological study on *Chlamydia trachomatis* infection among sexually active female adolescents. *Acta Obstet Gynaecol Jpn* 1990;**42**:1655–62.

179. Van Valkengoed IGM, Morre SA, van den Brule AJC, Meijer CJ, Bouter LM, van Eijk JT, *et al.* Follow-up, treatment, and reinfection rates among asymptomatic *Chlamydia trachomatis* cases in general practice. *Br J Gen Pract* 2002;**52**:623–7.

180. Morre SA, van den Brule AJC, Rozendaal L, Boeke AJ, Voorhorst FJ, deBlok S, *et al.* The natural course of asymptomatic *Chlamydia trachomatis* infections: 45% clearance and no development of clinical PID after one-year follow-up. *Int J STD AIDS* 2002;**13**:12–18. http://dx.doi.org/10.1258/095646202762226092

181. Sheffield JS, Andrews WW, Klebanoff MA, Macpherson C, Carey JC, Ernest JM, *et al.* Spontaneous resolution of asymptomatic *Chlamydia trachomatis* in pregnancy. *Obstet Gynecol* 2005;**105**:557–62. http://dx.doi.org/10.1097/01.AOG.0000153533.13658.c2

182. Alexander ER, Chandler J, Pheifer TA, Wang S-P, English M, Holmes KK. Prospective Study of Perinatal Chlamydia-Trachomatis Infection. In Hobson D, King Holmes, editors. *Nongonococcal Urethritis and Related Infections*. Washington DC; American Society for Microbiology; 1977.

183. Johannisson G, Lowhagen GB, Lycke E. Genital *Chlamydia trachomatis* infection in women. *Obstet Gynecol* 1980;**56**:671–5.

184. Paavonen J, Kousa M, Saikku P, Vartiainen E, Kanerva L, Lassus A. Treatment of non-Gonococcal urethritis with trimethoprim-sulfadiazine and with placebo: a double-blind partner-controlled study. *Brit J Vener Dis* 1980;**56**:101–4.

185. Sorensen JL, Thranov I, Hoff G, Dirach J. Early-onset and late-onset pelvic inflammatory disease among women with cervical *Chlamydia trachomatis* infection at the time of induced abortion: a follow-up study. *Infection* 1994;**22**:242–6. http://dx.doi.org/10.1007/BF01739907

186. McCormack WM, Alpert S, Mccomb DE, Nichols RL, Semine DZ, Zinner SH. 15-Month follow-up study of women infected with *Chlamydia trachomatis*. *N Engl J Med* 1979;**300**:123–5. http://dx.doi.org/10.1056/NEJM197901183000305

187. McCullagh P, Nelder JA. *Generalised Linear Models*. 2nd edn. London: Chapman & Hall; 1989. http://dx.doi.org/10.1007/978-1-4899-3242-6

188. Zelen M, Feinleib M. On the theory of screening for chronic diseases. *Biometrika* 1969;**56**:601–14. http://dx.doi.org/10.1093/biomet/56.3.601

189. Jamieson DJ, Theiler RN, Rasmussen SA. Emerging infections and pregnancy. *Emerg Infect Dis* 2006;**12**:1638–43. http://dx.doi.org/10.3201/eid1211.060152

190. Mallinson H, Hopwood J, Mutton K. Resolution of the recent performance problem of Abbott LCx *Chlamydia trachomatis* assay. Issues of repeat testing for confirmation of chlamydial infection. *Sex Transm Infect* 2002;**78**:225–6. http://dx.doi.org/10.1136/sti.78.3.225-a

191. Batteiger BE, Xu F, Johnson RE, Rekart ML. Protective immunity to *Chlamydia trachomatis* genital infection: evidence from human studies. *J Infect Dis* 2010;**202**: S178–89. http://dx.doi.org/10.1086/652400

192. LaMontagne DS, Baster K, Emmett L, Nicholls T, Randall S, McLean L, *et al.* Incidence and reinfection rates of genital chlamydial infection among women aged 16–24 years attending general practice, family planning and genitourinary medicine clinics in England: a prospective cohort study by the Chlamydia Recall Study Advisory Group. *Sex Transm Infect* 2007;**83**:292–303. http://dx.doi.org/10.1136/sti.2006.022053

193. Price M, Ades A, De Angelis D, Welton NJ, Mcleod J, Turner K, *et al.* Incidence of *Chlamydia trachomatis* infection in England: two methods of estimation. *Epidemiol Infect* 2014;**142**:562–67. http://dx.doi.org/10.1017/S0950268813001027

194. Geisler WM, Morrison SG, Bachmann LH. Absence of *Lymphogranuloma venereum* strains among rectal *Chlamydia trachomatis* outer membrane protein A genotypes infecting women and men who have sex with men in Birmingham, Alabama. *Sex Transm Dis* 2008;**35**:856–8. http://dx.doi.org/10.1097/OLQ.0b013e3181788313

195. Fenton KA, Korovessis C, Johnson AM, McCadden A, McManus S, Wellings K, *et al.* Sexual behaviour in Britain: reported sexually transmitted infections and prevalent genital *Chlamydia trachomatis* infection. *Lancet* 2001;**358**:1851–4. http://dx.doi.org/10.1016/S0140-6736(01)06886-6

196. McCadden A, Fenton KA, McManus S, Mercer CH, Erens B, Carder C, *et al. Chlamydia trachomatis* testing in the second British national survey of sexual attitudes and lifestyles: respondent uptake and treatment outcomes. *Sex Transm Dis* 2005;**32**:387–94. http://dx.doi.org/10.1097/01.olq.0000162364.65643.08

197. Office for National Statistics (ONS). *Population Estimates for Years 2000–2010*. London: ONS; 2014. (www.ons.gov.uk/ons/datasets-and-tables/index.html.)

198. Health Protection Agency (HPA). *STI Episodes seen in GUM Clinics: Annual Data Tables 2008*. London: HPA; 2010. (www.hpa.org.uk/stiannualdatatables.)

199. Rees E. The treatment of pelvic inflammatory disease. *Am J Obstet Gynecol* 1980;**138**:1042–7.

200. Hook EW, Spitters C, Reichart CA, Neumann TM, Quinn TC. Use of cell-culture and a rapid diagnostic assay for *Chlamydia trachomatis* screening. *JAMA* 1994;**272**:867–70. http://dx.doi.org/10.1001/jama.1994.03520110047027

201. Price M, Ades A, De Angelis D, Welton NJ, Mcleod J, Soldan K, *et al.* Risk of Pelvic Inflammatory Disease following *Chlamydia trachomatis* infection: analysis of prospective studies with a multistate model. *Am J Epidemiol* 2013;**178**:484–92. http://dx.doi.org/10.1093/aje/kws583

202. Bachmann LH, Richey CM, Waites K, Schwebke JR, Hook EW. Patterns of *Chlamydia trachomatis* testing and follow-up at a university hospital medical center. *Sex Transm Dis* 1999;**26**:496–9. http://dx.doi.org/10.1097/00007435-199910000-00002

203. Stamm WE, Guinan ME, Johnson C, Starcher T, Holmes KK, Mccormack WM. Effect of treatment regimens for *Neisseria gonorrhoeae* on simultaneous infection with *Chlamydia trachomatis*. *N Engl J Med* 1984;**310**:545–9. http://dx.doi.org/10.1056/NEJM198403013100901

204. Westergaard L, Philipsen T, Scheibel J. Significance of cervical *Chlamydia trachomatis* infection in post-abortal pelvic inflammatory disease. *Obstet Gynecol* 1982;**60**:322–5.

205. Michel C-EC, Sonnex C, Carne CA, White JA, Magbanua JP, Nadala JC Jr, *et al. Chlamydia trachomatis* load at matched anatomic sites: Implications for screening strategies. *J Clin Microbiol* 2007;**45**:1395–402. http://dx.doi.org/10.1128/JCM.00100-07

206. Geisler WM, Suchland RJ, Whittington WLH, Stamm WE. Quantitative culture of *Chlamydia trachomatis*: relationship of inclusion-forming units produced in culture to clinical manifestations and acute inflammation in urogenital disease. *J Infect Dis* 2001;**184**:1350–4. http://dx.doi.org/10.1086/323998

207. Wiggins R, Graf S, Low N, Horner PJ, Chlamydia Screening Studies (ClaSS) Screening Group. Real-time quantitative PCR to determine chlamydial load in men and women in a community setting. *J Clin Microbiol* 2009;**47**:1824–9. http://dx.doi.org/10.1128/JCM.00005-09

208. Herzog SA, Heijne JCM, Althaus CL, Low N. Describing the progression from *Chlamydia trachomatis* and *Neisseria gonorrhoeae* to pelvic inflammatory disease: systematic review of mathematical modeling studies. *Sex Transm Dis* 2012;**39**:628–37. http://dx.doi.org/10.1097/OLQ.0b013e31825159ff

209. Grimmett G, Stirzaker D. *Probability and Random Processes*. 2nd edn. Oxford: Oxford University Press; 1992.

210. Price MJ, Welton NJ, A.E A. Parameterisation of treatment effects for meta-analysis in multi-state Markov models. *Stat Med* 2011;**30**:140–51. http://dx.doi.org/10.1002/sim.4059

211. Welton NJ, Ades AE. Estimation of Markov Chain transition probabilities and rates from fully and partially observed data: uncertainty propagation, evidence synthesis and model calibration. *Med Decis Mak* 2005;**25**:633–45. http://dx.doi.org/10.1177/0272989X05282637

212. Taylor H, Karlin S. *An Introduction to Stochastic Modeling*. 3rd edn. San Diego, CA: Academic Press; 2006.

213. Boeke AJP, van Bergen JEAM, Morre SA, van Everdingen JJE. The risk of pelvic inflammatory disease associated with urogenital infection with *Chlamydia trachomatis*: literature review. *Ned Tijdschr Geneeskd* 2005;**149**:878–84.

214. Andersen B, Ostergaard L, Puho E, Skriver MV, Schonheyder HC. Ectopic pregnancies and reproductive capacity after *Chlamydia trachomatis* positive and negative test results: a historical follow-up study. *Sex Transm Dis* 2005;**32**:377–81. http://dx.doi.org/10.1097/01.olq.0000154512.86651.07

215. Bakken IJ, Ghaderi S. Incidence of pelvic inflammatory disease in a large cohort of women tested for *Chlamydia trachomatis*: a historical follow-up study. *BMC Infect Dis* 2009;**9**:130. http://dx.doi.org/10.1186/1471-2334-9-130

216. Bakken IJ, Skjeldestad FE, Lydersen S, Nordbo SA. Births and ectopic pregnancies in a large cohort of women tested for *Chlamydia trachomatis*. *Sex Transm Dis* 2007;**34**:739–43. http://dx.doi.org/10.1097/01.olq.0000261326.65503.f6

217. Bakken IJ, Skjeldestad FE, Nordbo SA. *Chlamydia trachomatis* infections increase the risk for ectopic pregnancy: A population-based, nested case–control study. *Sex Transm Dis* 2007;**34**:166–9. http://dx.doi.org/10.1097/01.olq.0000230428.06837.f7

218. Brihmer C, Kallings I, Nord CE, Brundin J. Salpingitis: aspects of diagnosis and etiology – a 4-year study from a Swedish capital hospital. *Eur J Obstet Gynecol Reprod Biol* 1987;**24**:211–20. http://dx.doi.org/10.1016/0028-2243(87)90020-7

219. Aghaizu A, Adams EJ, Turner K, Kerry S, Hay P, Simms I, *et al.* What is the cost of pelvic inflammatory disease and how much could be prevented by screening for *Chlamydia trachomatis*? Cost analysis of the Prevention Of Pelvic Infection (POPI) trial. *Sex Transm Infect* 2011;**87**:312–17. http://dx.doi.org/10.1136/sti.2010.048694

220. Mardh PA. Tubal factor infertility, with special regard to chlamydial salpingitis. *Curr Opin Infect Dis* 2004;**17**:49–52. http://dx.doi.org/10.1097/00001432-200402000-00010

221. Oakeshott P, Kerry S, Atherton H, Aghaizu A, Hay S, Taylor Robinson D, *et al.* Community-based trial of screening for *Chlamydia trachomatis* to prevent pelvic inflammatory disease: the POPI (prevention of pelvic infection) trial. *Trials* 2008;**9**:73. http://dx.doi.org/10.1186/1745-6215-9-73

222. Herzog S, Althaus C, Heijne J, Oakeshott P, Kerry S, Hay P, *et al.* Timing of progression from *Chlamydia trachomatis* infection to pelvic inflammatory disease: a mathematical modelling study. *BMC Infect Dis* 2012;**12**:187. http://dx.doi.org/10.1186/1471-2334-12-187

223. Lazaro N. *Sexually Transmitted Infections in Primary Care*. 2006. RCGP/BASHH. London: BMJ Publishing Group; 2006. URL: www.rcgp.org and www.bashh.org/guidelines

224. Bevan CD, Johal BJ, Mumtaz G, Ridgway GL, Siddle NC. Clinical, laparoscopic and microbiological findings in acute salpingitis: report on a United Kingdom cohort. *Br J Obstet Gynaecol* 1995;**102**:407–14. http://dx.doi.org/10.1111/j.1471-0528.1995.tb11294.x

225. Hoyme UB, Kentner A, Mylonas I. Laparoscopic diagnosis of chlamydial pelvic inflammatory disease and its impact on chlamydia screening programmes. *Eur Obstet Gynaecol* 2012;**7**:9–13.

226. Gottlieb SL, Xu F, Brunham RC. Screening and treating *Chlamydia trachomatis* genital infection to prevent pelvic inflammatory disease: interpretation of findings from randomized controlled trials. *Sex Transm Dis* 2013;**40**:97–102. http://dx.doi.org/10.1097/olq.0b013e31827bd637

227. Wendel GD, Cox SM, Bawdon RE, Theriot SK, Heard MC, Nobles BJ. A randomized trial of ofloxacin versus cefoxitin and doxycycline in the outpatient treatment of acute salpingitis. *Am J Obstet Gynecol* 1991;**164**:1390–6. http://dx.doi.org/10.1016/0002-9378(91)91479-G

228. Ness RB, Soper DE, Holley RL, Peipert J, Randall H, Sweet RL, *et al.* Effectiveness of inpatient and outpatient treatment strategies for women with pelvic inflammatory disease: results from the Pelvic Inflammatory Disease Evaluation and Clinical Health (PEACH) Randomized Trial. *Am J Obstet Gynecol* 2002;**186**:929–37. http://dx.doi.org/10.1067/mob.2002.121625

229. Simms I, Eastick K, Mallinson H, Thomas K, Gokhale R, Hay P, *et al.* Associations between *Mycoplasma genitalium*, *Chlamydia trachomatis* and pelvic inflammatory disease. *J Clin Pathol* 2003;**56**:616–18. http://dx.doi.org/10.1136/jcp.56.8.616

230. Mascellino MT, Ciardi MR, Oliva A, Cecinato F, Hassemer MP, Borgese L. *Chlamydia trachomatis* detection in a population of asymptomatic and symptomatic women: correlation with the presence of serological markers for this infection. *New Microbiol* 2008;**31**:249–56.

231. Hughes G, Williams T, Simms I. *Trends in Diagnoses of Genital Chlamydia and Pelvic Inflammatory Disease in Primary Care, 1990–2003: Analysis of Data in the General Practice Research Database*. Health Protection Agency Annual Conference, University of Warwick, Warwick, 13–15 September 2004.

232. Westrom L. Incidence, prevalence, and trends of acute pelvic inflammatory disease and its consequences in industrialized countries. *Am J Obstet Gynecol* 1980;**138**:880–92.

233. Simms I, Warburton F, Westrom L. Diagnosis of pelvic inflammatory disease: time for a rethink. *Sex Transm Infect* 2003;**79**:491–4. http://dx.doi.org/10.1136/sti.79.6.491

234. Ross J, McCarthy G. *UK National Guideline for the Management of Pelvic Inflammatory Disease*. Macclesfield: BASHH; 2011.

235. Taylor-Robinson D, Jensen JS, Svenstrup HF, Stacey CM. Difficulties experienced in defining the microbial cause of pelvic inflammatory disease. *Int J STD AIDS* 2012;**23**:18–24. http://dx.doi.org/10.1258/ijsa.2011.011066

236. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS: a Bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput* 2000;**10**:325–37. http://dx.doi.org/10.1023/A:1008929526011

237. Bjartling C, Osser S, Persson K. *Mycoplasma genitalium* in cervicitis and pelvic inflammatory disease among women at a gynecologic outpatient service. *Am J Obstet Gynecol* 2012;**206**:476.e1–8. http://dx.doi.org/10.1016/j.ajog.2012.02.036

238. Chandra A, Martinez GM, Mosher WD, Abma JC, Jones J. Fertility, family planning, and reproductive health of U.S. women: data from the 2002 National Survey of Family Growth. *NVHS Vital Health Stat* 2005;**23**:1–160.

239. Buchan H, Vessey M, Goldacre M, Fairweather J. Morbidity following pelvic inflammatory disease. *Br J Obstet Gynaecol* 1993;**100**:558–62. http://dx.doi.org/10.1111/j.1471-0528.1993.tb15308.x

240. Heinonen PK, Leinonen M. Fecundity and morbidity following acute pelvic inflammatory disease treated with doxycycline and metronidazole. *Arch Gynecol Obstet* 2003;**268**:284–8. http://dx.doi.org/10.1007/s00404-002-0376-6

241. Safrin S, Schachter J, Dahrouge D, Sweet RL. Long-term sequelae of acute pelvic inflammatory disease: a retrospective cohort study. *Am J Obstet Gynecol* 1992;**166**:1300–5. http://dx.doi.org/10.1016/S0002-9378(11)90626-7

242. Gerber B, Krause A. A study of second-look laparoscopy after acute salpingitis. *Arch Gynecol Obstet* 1996;**258**:193–200. http://dx.doi.org/10.1007/s004040050123

243. Bernsteine R, Kennedy WR, Waldron J. Acute pelvic inflammatory disease: a clinical follow-up. *Int J Fertil* 1987;**32**:229–32.

244. Grimes DA. Intrauterine device and upper-genital-tract infection. *Lancet* 2000;**356**:1013–19. http://dx.doi.org/10.1016/S0140-6736(00)02699-4

245. Farley TMM, Rowe PJ, Meirik O, Rosenberg MJ, Chen JH. Intrauterine devices and pelvic inflammatory disease: an international perspective. *Lancet* 1992;**339**:785–8. http://dx.doi.org/10.1016/0140-6736(92)91904-M

246. Bouyer J, Coste J, Shojaei T, Pouly JL, Fernandez H, Gerbaud L, *et al.* Risk factors for ectopic pregnancy: a comprehensive analysis based on a large case-control, population-based study in France. *Am J Epidemiol* 2011;**157**:185–94. http://dx.doi.org/10.1093/aje/kwf190

247. Coste J, Laumon B, Bremond A, Collet P, Jobspira N. Sexually-transmitted diseases as major causes of ectopic pregnancy: results from a large case–control study in France. *Fertil Steril* 1994;**62**:289–95.

248. Ramsey SD, Sullivan SD. Weighing the economic evidence: guidelines for the critical assessment of cost-effectiveness analyses. *J Am Board Fam Pract* 1999;**12**:477–85. http://dx.doi.org/10.3122/jabfm.12.6.477

249. Dominici F. Combining contingency tables with missing dimensions. *Biometrics* 2000;**56**:546–53. http://dx.doi.org/10.1111/j.0006-341X.2000.00546.x

250. Dominici F, Parmigiani G, Reckhow KH, Wolpert RL. Combining information from related regressions. *JABES* 1997;**2**:313–32. http://dx.doi.org/10.2307/1400448

251. Teisala K, Heinonen PK, Aine R, Punnonen R, Paavonen J. Second laparoscopy after treatment of acute pelvic inflammatory disease. *Obstet Gynecol* 1987;**69**:343–6.

252. Wolner-Hanssen P, Westrom L. Second look laparoscopy after acute salpingitis. *Obstet Gynecol* 1983;**61**:702–4.

253. Pavletic AJ, Wolner-Hanssen P, Paavonen J, Hawes SE, Eschenbach DA. Infertility following pelvic inflammatory disease. *Infect Dis Obstet Gynecol* 1999;**7**:145–52. http://dx.doi.org/10.1155/S106474499900023X

254. Brihmer C, Kallings I, Nord CE, Brundin J. Second look laparoscopy; evaluation of 2 different antibiotic regimens after treatment of acute salpingitis. *Eur J Obstet Gynecol Reprod Biol* 1989;**30**:263–74. http://dx.doi.org/10.1016/0028-2243(89)90011-7

255. Maheshwari A, Hamilton M, Bhattacharya S. Effect of female age on the diagnostic categories of infertility. *Hum Reprod* 2008;**23**:538–42. http://dx.doi.org/10.1093/humrep/dem431

256. Kavanagh K, Wallace LA, Robertson C, Wilson P, Scoular A. Estimation of the risk of tubal factor infertility associated with genital chlamydial infection in women: a statistical modelling study. *Int J Epidemiol* 2013;**42**:493–503. http://dx.doi.org/10.1093/ije/dyt011

257. Hull MGR, Glazener CMA, Kelly NJ, Conway DI, Foster PA, Hinton RA, *et al.* Population study of causes, treatment, and outcome of infertility. *BMJ* 1985;**291**:1693–7. http://dx.doi.org/10.1136/bmj.291.6510.1693

258. Sweet RL. Pelvic inflammatory disease: current concepts of diagnosis and management. *Curr Infect Dis Rep* 2012;**14**:194–203. http://dx.doi.org/10.1007/s11908-012-0243-y

259. Ness RB, Soper DE, Richter HE, Randall H, Peipert JF, Nelson DB, *et al.* Chlamydia antibodies, chlamydia heat shock protein, and adverse sequelae after pelvic inflammatory disease: The PID Evaluation and Clinical Health (PEACH) Study. *Sex Transm Dis* 2008;**35**:129–35. http://dx.doi.org/10.1097/OLQ.0b013e3181557c25

260. Ross JDC, Radcliffe K, Ahmed-Jushuf I, Cowan F, Fitzgerald M, Wilson J, *et al.* National guideline for the management of pelvic infection and perihepatitis. *Sex Transm Infect* 1999;**75**:S54–6.

261. Cates W, Wasserheit JN. Genital Chlamydial infections: epidemiology and reproductive sequelae. *Am J Obstet Gynecol* 1991;**164**:1771–81. http://dx.doi.org/10.1016/0002-9378(91)90559-A

262. Persson K. The role of serology, antibiotic susceptibility testing and serovar determination in genital chlamydial infections. *Best Pract Res Clin Obstet Gynecol* 2002;**16**:801–14. http://dx.doi.org/10.1053/beog.2002.0321

263. Kosseim M, Brunham RC. Fallopian tube obstruction as a sequela to *Chlamydia trachomatis* infection. *European J Clin Microbiol* 1986;**5**:584–90. http://dx.doi.org/10.1007/BF02017711

264. Johnson AM, Horner P. A new role for *Chlamydia trachomatis* serology? *Sex Transm Infect* 2008;**84**:79–80. http://dx.doi.org/10.1136/sti.2007.028472

265. Morré SA, Munk C, Persson K, Krüger-Kjaer S, van Dijk R, Meijer CJ, *et al.* Comparison of three commercially available peptide-based immunoglobulin G (IgG) and IgA assays to microimmunofluorescence assay for detection of *Chlamydia trachomatis* antibodies. *J Clin Microbiol* 2002;**40**:584–7. http://dx.doi.org/10.1128/JCM.40.2.584-587.2002

266. Hadgu A. Discrepant analysis: a biased and an unscientific method for estimating test sensitivity and specificity. *J Clin Epidemiol* 1999;**52**:1231–7. http://dx.doi.org/10.1016/S0895-4356(99)00101-8

267. Miller WC. Bias in discrepant analysis: when two wrongs don't make a right. *J Clin Epidemiol* 1998;**51**:219–31. http://dx.doi.org/10.1016/S0895-4356(97)00264-3

268. Narvanen A, Puolakkainen M, Hao W, Kino K, Suni J. Detection of antibodies to *Chlamydia trachomatis* with peptide-based species-specific enzyme immunoassay. *Infect Dis Obstet Gynecol* 1997;**5**.

269. Wills GS, Horner PJ, Reynolds R, Johnson AM, Muir DA, Brown DW, *et al.* Pgp3 antibody enzyme-linked immunosorbent assay, a sensitive and specific assay for seroepidemiological analysis of *Chlamydia trachomatis* Infection. *Clin Vaccine Immunol* 2009;**16**:835–43. http://dx.doi.org/10.1128/CVI.00021-09

270. Akande VA, Hunt LP, Cahill DJ, Caul EO, Ford WCL, Jenkins JM. Tubal damage in infertile women: prediction using chlamydia serology. *Hum Reprod* 2003;**18**:1841–7. http://dx.doi.org/10.1093/humrep/deg347

271. Conway D, Glazener CM, Caul EO, Hodgson J, Hull MG, Clarke SK, *et al.* Chlamydial serology in fertile and infertile women. *Lancet* 1984;**1**:191–3. http://dx.doi.org/10.1016/S0140-6736(84)92113-5

272. Darville T, Hiltke TJ. Pathogenesis of genital tract disease due to *Chlamydia trachomatis*. *J Infect Dis* 2010;**201**: S114–25. http://dx.doi.org/10.1086/652397

273. Kihlstrom E, Lindgren R, Ryden G. Antibodies to *Chlamydia trachomatis* in women with infertility, pelvic inflammatory disease and ectopic pregnancy. *Eur J Obstet Gynecol Reprod Biol* 1990;**35**:199–204. http://dx.doi.org/10.1016/0028-2243(90)90162-T

274. Land JA, Gijsen AP, Kessels AGH, Slobbe MEP, Bruggeman CA. Performance of five serological chlamydia antibody tests in subfertile women. *Hum Reprod* 2003;**18**:2621–7. http://dx.doi.org/10.1093/humrep/deg479

275. Price MJ, Ades AE, Welton NJ, Macleod J, Turner K, Simms I, *et al.* How much tubal factor infertility is caused by chlamydia? Estimation based on serological evidence adjusted for sensitivity and specificity. *J Sex Transm Dis* 2012;**39**:608–13. http://dx.doi.org/10.1097/OLQ.0b013e3182572475

276. Broeze KA, Opmeer BC, Coppus SF, Van Geloven N, Ales MF, Anestad G, *et al.* Chlamydia antibody testing and diagnosing tubal pathology in subfertile women: an individual patient data meta-analysis. *Hum Reprod Update* 2011;**17**:301–10. http://dx.doi.org/10.1093/humupd/dmq060

277. Green DM, Swets JA. *Signal Detection Theory and Psychophysics*. New York, NY: Wiley; 1966.

278. Persson K, Osser S, Birkelund S, Christiansen G, Brade H. Antibodies to *Chlamydia trachomatis* heat shock proteins in women with tubal factor infertility are associated with prior infection by *C. trachomatis* but not by *C. pneumoniae*. *Hum Reprod* 1999;**14**:1969–73. http://dx.doi.org/10.1093/humrep/14.8.1969

279. Claeskens G, Hjort NL. *Model Selection and Model Averaging*. Cambridge: Cambridge University Press; 2011.

280. Horner P, Wills G, Pebody R. A Comparative Study of *Chlamydia trachomatis* PGP3 Antibody Elisa and Two *C. trachomatis* Specific MOMP Peptide Assays on Sera from Young People in the Community. In Stary A, Byrne G, Caldwell HD, editors. Twelfth International Symposium on Human Chlamydial Infections, San Francisco, CA, 2010. pp. 453–6.

281. Horner PJ, Wills GS, Reynolds R, Johnson AM, Muir DA, Winston A, *et al.* Effect of time since exposure to *Chlamydia trachomatis* on chlamydia antibody detection in women: a cross-sectional study. *Sex Transm Infect* 2013;**89**:398–403. http://dx.doi.org/10.1136/sextrans-2011-050386

282. Chavkin W. The rise in ectopic pregnancy: exploration of possible reasons. *Int J Gynecol Obstet* 1982;**20**:341–50. http://dx.doi.org/10.1016/0020-7292(82)90066-2

283. Westrom L. Effect of acute pelvic inflammatory disease on fertility. *Am J Obstet Gynecol* 1975;**121**:707–13. http://dx.doi.org/10.1097/00006254-197508000-00022

284. Turner K, Adams E, Grant A, Macleod J, Bell G, Clarke J, *et al.* Costs and cost effectiveness of different strategies for chlamydia screening and partner notification: an economic and mathematical modelling study. *BMJ* 2011;**342**:c7250. http://dx.doi.org/10.1136/bmj.c7250

285. Anestad G, Lunde O, Moen M, Dalaker K. Infertility and chlamydial infection. *Fertil Steril* 1987;**48**:787–90.

286. McLachlan GJ, Peel D. *Finite Mixture Models*. New York, NY: Wiley; 2000. http://dx.doi.org/10.1002/0471721182

287. Soper DE. Pelvic inflammatory disease. *Obstet Gynecol* 2010;**116**:1459–60. http://dx.doi.org/10.1097/AOG.0b013e3181ffb0de

288. Garnett GP, Cousens S, Hallett TB, Steketee R, Walker N. Mathematical models in the evaluation of health programmes. *Lancet* 2011;**378**:515–25. http://dx.doi.org/10.1016/S0140-6736(10)61505-X

289. Presanis AM, De Angelis D, Goubar A, Gill ON, Ades AE. Bayesian evidence synthesis for a transmission dynamic model for HIV among men who have sex with men. *Biostatistics* 2011;**12**:666–81. http://dx.doi.org/10.1093/biostatistics/kxr006

290. Birrell P, Ketsetzis G, Gay N, Cooper BS, Presanis AM, Harris RJ, *et al.* Bayesian modeling to unmask and predict influenza A/H1N1pdm dynamics in London. *Proc Nat Acad Sci U S A* 2011;**108**:18238–43. http://dx.doi.org/10.1073/pnas.1103002108

291. De Angelis D, Presanis A, Birrell P, Tomba G, House T. Four key challenges in infectious disease modelling using data from multiple sources. *Epidemics* 2015;**10**:83–7. http://dx.doi.org/10.1016/j.epidem.2014.09.004

292. Van Valkengoed GM. *Asymptomatic Chlamydia trachomatis Infections: Should We Screen?* Doctoral Dissertation. Wageningen: Vrije Universiteit, Ponsen en Looijen BV; 2001.

293. Dunson D, Baird D, Colombo B. Increased infertility with age in men and women. *Obstet Gynecol* 2004;**103**:51–6. http://dx.doi.org/10.1097/01.AOG.0000100153.24061.45

Appendix 1 Mean duration of follow-up in the *Chlamydia trachomatis* duration studies (see *Chapter 4*)

This appendix gives details of how mean duration of follow-up was estimated for the studies in *Table 4* when it was not specified in the original publication.

- Johannisson *et al.*[183] report a group of patients followed up after 5–8 weeks and we assume a mean follow-up of $(5 + 8)/2 = 6.5$ weeks.
- For the study by Joyner *et al.*,[79] the centres of the reported ranges are used except for the last observation were we use the geometric mean = 100 days.
- Geisler *et al.*[73] report a median follow-up of 13 days, with a range of 4–59 days. The distribution is skewed but is approximately symmetrical on the log scale. If μ and σ are, respectively, the mean and SD of the natural log of the follow-up time then the mean follow-up is $\exp(\mu + \sigma^2/2)$. We estimate σ as being one-quarter of the range and μ is the natural log of the reported median follow-up.
- Sorensen *et al.*[185] report a median of 11 months and a range of 2–24 months. This distribution is nearly symmetrical. The mean is estimated as $[11 + (24 + 2)/2]/2 = 12$ months.
- McCormack *et al.*[186] reported a follow-up of 16–17 months, so we assume a mean follow-up of 16.5 months.
- Sheffield *et al.*[181] report that patients were examined 16–23 weeks into pregnancy and a second time at 24–29 weeks into pregnancy. In total, 52 out of 140 were followed up for < 5 weeks. The minimum follow-up is therefore $24 - 23 = 1$ week and the maximum follow-up is $29 - 16 = 13$ weeks. We assume that the first group were followed for 3 weeks, and the second group for 9 weeks.

Appendix 2 WinBUGS code and data for the two-rate model (see *Chapter 4*)

WinBUGS code and Data for the 2-rate model

```
model {
# likelihood
for (i in 1:studnum) {
 for (j in 1:studobs[i]) {
  r[i,j] ~ dbin(theta[i,j],n[i,j])
  }
 }

# model
for (i in 1:studnum-2) {
 for (j in 1:studobs[i]) {
  theta[i,j] <- ((z[i,j] / n[i,j]) * (1 - exp(-lambda.C[1] * t[i,j]))
+                    (1 - (z[i,j] / n[i,j])) * (1 - exp(-lambda.C[2]
* t[i,j])))
                / (psi / (1 - ((1-psi) * equals(seind[i,j],0))))
  }
 }
# Left-truncated studies with repeat observations
for (i in 8:9) {
 for (j in 1:studobs[i]) {
  temp[i,j,1] <- ((z[i,1] / n[i,1]) * exp(-lambda.C[1] * T[i,j]) /
                 ((z[j,1] / n[i,1]) * exp(-lambda.C[1] * T[i,j]) +
                 (1 - (z[i,1] / n[i,1])) * exp(-lambda.C[2] *
T[i,j]))) *
                 (1 - exp(-lambda.C[1] * t[i,j]))
  temp[i,j,2] <- ((1 - (z[i,j] / n[i,1])) * exp(-lambda.C[2] *
T[i,j]) /
                  (0.00001+(z[i,1] / n[i,1]) * exp(-lambda.C[1] *
T[i,j]) +
                  (1 - (z[i,1] / n[i,1])) * exp(-lambda.C[2] *
T[i,j]))) *
                  (1 - exp(-lambda.C[2] * t[i,j]))
  theta[i,j] <- (temp[i,j,1] + temp[i,j,2]) /
                (psi / (1 - ((1-psi) * equals(seind[i,j],0))))
  }
 }

# priors
p1 ~ dbeta(1,1)
lambda.C[1] <- 120
lambda.C[2] ~ dexp(0.001)
psi ~ dbeta(78,8)  #sensitivity of culture given initial positive
# culture

# Class proportions
# t=0 studies
for (i in 1:4) {
 for (j in 1:studobs[i]) {
  z[i,j] ~ dbin(p1,n[i,j]) # start at t=0
  }
 }

# Left-truncated studys
for (i in 5:studnum) {
 for (j in 1:studobs[i]) {
  z[i,j] ~ dbin(w1,n[i,j])
  }
```

```
 }

# deviance
for (i in 1:studnum) {
 for (j in 1:studobs[i]) {
   dev[i,j] <- 2 * (r[i,j] * log(r[i,j] / (theta[i,j] * n[i,j])) +
                    (n[i,j] - r[i,j]) * log((n[i,j] - r[i,j]) /
                    (n[i,j] - (n[i,j] * theta[i,j]))))
   }
 dev.stud[i] <- sum(dev[i,1:studobs[i]])
 }
sumdev <- sum(dev.stud[])

# left truncation
w1 <- (p1 / lambda.C[1]) / (p1 / lambda.C[1] + (1 - p1) /
lambda.C[2])

# summary statistics
dur <- 1 / lambda.C[2]

# Predicted values for Forest plot
for (i in 1:studnum) {
 for (j in 1:studobs[i]) {
  stud.lambda.Cexpect[i,j] <- -log(1 - theta[i,j]) / t[i,j]
  stud.dur.expect[i,j] <- 1 / stud.lambda.Cexpect[i,j]
  }
 }
}

# Data
list(

# duration
# study order
# 1 Johhanisson
# 2 Joyner
# 3 Geisler
# 4 Paavonen
# 5 Rahm
# 6 Sorensen
# 7 McCormack
# 8 Morre
# 9 Mollano

r = structure(.Data=c(
10,7,6,6,NA,
2,7,1,0,3,
23,NA,NA,NA,NA,
3,NA,NA,NA,NA,
17,0,0,NA,NA,
8,NA,NA,NA,NA,
3,NA,NA,NA,NA,
2,2,4,0,2,
44,23,7,2,NA
),.Dim=c(9,5)),

n = structure(.Data=c(
23,14,14,8,NA,
12,28,4,8,6,
129,NA,NA,NA,NA,
15,NA,NA,NA,NA,
```

```
93,1,1,NA,NA,
13,NA,NA,NA,NA,
7,NA,NA,NA,NA,
20,5,15,1,13,
82,37,14,6,NA
),.Dim=c(9,5)),

t=structure(.Data=c(
0.038,0.058,0.077,0.125,NA,
0.012,0.03,0.049,0.088,0.274,
0.045,NA,NA,NA,NA,
0.083,NA,NA,NA,NA,
0.25,0.5,0.75,NA,NA,
1,NA,NA,NA,NA,
1.375,NA,NA,NA,NA,
0.083,0.5,0.417,0.917,0.5,
1,1,1,1,NA
),.Dim=c(9,5)),

seind = structure(.Data=c(
1,1,1,1,NA,
0,0,0,0,0,
0,NA,NA,NA,NA,
1,NA,NA,NA,NA,
1,1,1,NA,NA,
0,NA,NA,NA,NA,
1,NA,NA,NA,NA,
0,0,0,0,0,
0,0,0,0,NA
),.Dim=c(9,5)),

T=structure(.Data=c(
NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,
0,0,0.083,0.083,0.5,
0,1,2,3,NA
),.Dim=c(9,5)),

studnum = 9,
studobs = c(4,5,1,1,3,1,1,5,4),
)

# Initial values - 1
list(
psi = 0.9,
lambda.C = c(NA,0.7),
p1 = 0.2,
)

# Initial values - 2
list(
psi = 0.6,
lambda.C = c(NA,0.1),
p1 = 0.5
)
```

Appendix 3 Assessment of statistical modelling assumptions in *Chapter 5*

Appendix 5.1: Regression analysis of the LaMontagne data

We fit the following nine regression models to the 18 data points from the LaMontagne study[192] shown in *Table 9*:

- Model 1 $\log(\lambda_{a,s,i}) = \alpha + \gamma_a + \rho_s + \eta_i + \gamma\rho_{as} + \gamma\eta_{ai} + \rho\eta_{si} + \gamma\rho\eta_{asi}$
- Model 2 $\log(\lambda_{a,s,i}) = \alpha + \gamma_a + \rho_s + \eta_i + \gamma\rho_{as} + \gamma\eta_{ai} + \rho\eta_{si}$
- Model 3 $\log(\lambda_{a,s,i}) = \alpha + \gamma_a + \rho_s + \eta_i + \gamma\eta_{ai} + \rho\eta_{si}$
- Model 4 $\log(\lambda_{a,s,i}) = \alpha + \gamma_a + \rho_s + \eta_i + \gamma\rho_{as} + \rho\eta_{si}$
- Model 5 $\log(\lambda_{a,s,i}) = \alpha + \gamma_a + \rho_s + \eta_i + \gamma\rho_{as} + \gamma\eta_{ai}$
- Model 6 $\log(\lambda_{a,s,i}) = \alpha + \gamma_a + \rho_s + \eta_i + \gamma\rho_{as}$
- Model 7 $\log(\lambda_{a,s,i}) = \alpha + \gamma_a + \rho_s + \eta_i + \gamma\eta_{ai}$
- Model 8 $\log(\lambda_{a,s,i}) = \alpha + \gamma_a + \rho_s + \eta_i + \rho\eta_{si}$
- Model 9 $\log(\lambda_{a,s,i}) = \alpha + \gamma_a + \rho_s + \eta_i$,

where γ_1, ρ_1, η_1, $\gamma\rho_{a1}$, $\gamma\rho_{1s}$, $\gamma\eta_{ai}$, $\gamma\eta_{1i}$, $\rho\eta_{s1}$, $\rho\eta_{l1}$, $\gamma\rho\eta_{as1}$, $\gamma\rho\eta_{a1i}$, $\gamma\rho\eta_{1si} = 0$; $\gamma\rho_{as}$ represents an interaction between age and setting, $\gamma\eta_{ai}$ between age and reinfection, $\rho\eta_{si}$ between setting and reinfection, and $\gamma\rho\eta_{asi}$ is a three-way interaction between age, setting and re-infection.

The estimates of $\lambda_{a,s,i}$ along with estimates of duration are used to estimate $\kappa(t)_{a,s,i}$, which is the parameter in the likelihood function for the LaMontagne data[192] (see *Figure 8*, DAG). Model fit statistics together with the nominal numbers of parameters for each model are shown in *Table 43*. Results are based on two chains, run for 40,000 samples after a 10,000 burn-in. The results show that Model 8, which includes only the main effects and an interaction between setting and infection/re-infection status has the lowest DIC (see *Chapter 3*). A plot of the deviance residuals for Model 8 revealed no evidence of a systematic bad fit. Model 8 is identical to the one described in the main text, although it has been re-parameterised slightly to simplify the notation. It is only marginally better in terms of DIC than a model that also includes an interaction between age and setting (Model 4) or a model that assumes no interactions (Model 9).

TABLE 43 Model fit statistics for each of the regression models fitted to infection and re-infection data, *Chapter 5*

Model	Residual deviance	Nominal no. of parameters	DIC	p_D
1	18.6	18	103.2	17.9
2	15.8	14	96.2	13.2
3	19.0	10	95.8	10.1
4	14.5	12	93.1	11.9
5	22.6	12	101.1	11.9
6	21.1	10	97.7	10.0
7	24.1	8	98.9	8.0
8	17.3	8	91.9	8.0
9	22.3	6	95.0	6.0

Appendix 5.2: assumed relationship between setting-specific odds ratios from Adams and hazard ratios in the model for the LaMontagne data

We assume that the between-setting ORs are equivalent to between-setting RRs because of the rare disease assumption, and that these inform the between-setting HRs in the LaMontagne study.[192] This is not strictly correct, as they should inform the between-setting incidence ratios.

We assess the sensitivity of the results to this assumption. *Table 44* shows the between-setting infection ratios (column 1), FOI ratios (column 2) and incidence ratios (column 3) estimated from the LaMontagne data[192] alone. The corresponding results from Adams *et al.*[106] (introduced in *Table 11*) are repeated in column 4. The infection ratios from LaMontagne *et al.*[192] are almost identical to the ORs from Adams *et al.*[106] However, this is not a reason to conclude that our model is better than the 'correct' model, where they inform incidence ratios. There is some discrepancy between the infection ratios compared with the incidence or FOI ratios. The FOI and incidence ratios are almost identical.

Because of the lack of data for the general population in LaMontagne *et al.*,[192] it is very difficult to correctly parameterise the model so that the ORs inform the incidence ratios. It is, however, possible, although considerably more mathematically complicated than the model described in *Chapter 5*, to apply the ORs to FOI ratios. We performed this analysis for the full synthesis model and found that incidence changed by less than a multiplicative factor of 5% in all age groups (mean < 3%). From this, and the fact that the incidence and FO ratios agree so closely, we conclude that there is only very negligible bias from not parameterising the model so that ORs inform incidence ratios.

TABLE 44 Between-setting ratios of infection, FOI and incidence rates, *Chapter 5*

	LaMontagne (2007)[192]			Adams (2004)[106]
Ratio	Infection ratios (95% CrI)	FOI ratios (95% CrI)	Incidence ratios (95% CrI)	OR (95% CI) (repeated from *Table 11*)
FP to GP	1.27 (0.59 to 2.48)	1.10 (0.65 to 1.74)	1.11 (0.61 to 1.88)	1.27 (1.00 to 1.62)
STD to GP	2.31 (1.11 to 4.45)	1.60 (0.95 to 2.53)	1.74 (0.94 to 2.97)	2.39 (0.72 to 3.33)

Appendix 4 Details of the statistical methods in *Chapter 5*

A full list of basic and functional parameters along with brief descriptions is provided in *Table 45*, which should be viewed alongside the DAG (see *Figure 8*).

TABLE 45 Master list of parameters for *Chapter 5*

Basic parameters	Priors	Interpretation
Incidence		
$\lambda_{1,1,1}$	Log-normal	CT infection rate in women aged 16–17 years in the GP setting
γ_a	Log-normal	HR across age groups
ρ_s	Log-normal	HR across settings
η_s	Log-normal	Setting specific re-infection to infection HRs
$p_{a,s}$	Beta	Baseline prevalence of LaMontagne[192] sampling frame
Duration		
Δ_s	Normal	Mean duration of symptomatic CT infection
Δ_A	Uniform	Mean duration of asymptomatic CT infection
ϕ	Binomial	Proportion of CT episodes in which symptoms will develop
Prevalence		
$\pi_{a.pop}$ *Note: Prevalence is a basic parameter in method B but a functional parameter in the full synthesis model*	$\pi_{a.pop} = \pi_{1.pop}.\gamma_a$	Population prevalence of CT by age group
Functional parameters	**Function**	**Interpretation**
Incidence		
$\lambda_{a,s,1}$	$\gamma_a.\rho.\lambda_{111}$	Infection rate for women in age group *a* and setting *s*
$\lambda_{a,s,2}$	$\eta_s.\lambda_{a,s,1}$	Re-infection rate for women in age group *a* and setting *s*
$\tilde{\lambda}^{FOI}_{a,GP}$	$(1-p_{a,s}).\lambda_{a,s,1} + p_{a,s}.\lambda_{a,s,2}$	FOI for women in age group a in the GP setting
$\tilde{\lambda}^{FOI}_{a,pop}$	$\rho_{pop}\tilde{\lambda}^{FO}_{a,GP}$	FOI for women in age group a in the general population
$\tilde{\lambda}^{INC1}_{a,pop}$	$\frac{\lambda^{FO}_{a,pop}}{1-\lambda^{FO}_{a,pop}\Delta}$	Incidence rate of CT for women in age group *a* in the general population estimated using method A
$\tilde{\lambda}^{INC2}_{a,pop}$	$\frac{\pi_{a,pop}}{\Delta}$	Incidence rate of CT for women in age group a in the general population estimated using method B
$\kappa(t)_{a,s,i}$	See *Figure 8*	Proportion of CT– women in age group *a*, setting *s* and re-infection status *i*, expected to be CT+ after 6 months

continued

TABLE 45 Master list of parameters for *Chapter 5* (*continued*)

Basic parameters		Priors	Interpretation
Duration			
Δ		$\Delta_S.\phi + \Delta_A.(1-\phi)$	Mean duration of CT infection
λ_A^C		$1/\Delta_A$	Mean clearance rate of asymptomatic CT infections
λ_S^C		$1/\Delta_S$	Mean clearance rate of symptomatic CT infections
Prevalence			
$\pi_{a,pop}$	*Note: Prevalence is a basic parameter in method B but a functional parameter in the full synthesis model*	$\tilde{\lambda}^{INC}_{a,pop}.\Delta$	Population prevalence of CT by age group

Most of the functional relationships have been spelled out in the methods section as equations 11–17, or in the DAG. Some expressions require further explanation.

The expression:

$$\kappa(t) = 1 - \frac{\lambda^C + \lambda e^{-(\lambda+\lambda^C)t}}{\lambda + \lambda^C} \tag{73}$$

relates the proportion of infected individuals $\kappa(t)$, who were initially uninfected, observed after time t, to an incidence rate λ and a clearance rate λ^C. This can be derived from the Kolmogorov's forward equations.[209] In the DAG, the more complex relationship for proportion infected at time t represents a weighted average of two clearance rates: one being in symptomatic women and the other in asymptomatic women. The proportion of infections in which symptoms develop is the weight, and the clearance rates in each group are the reciprocals of the mean duration of symptomatic and asymptomatic infections.

Strictly speaking, *Figure 8* sets out the relationships as they would be if the incidence and prevalence data were available on the exact same age groups. As the age groupings in the studies were slightly different, we used Census information on the English female population sizes from 2002 for each year of age 16–44 years to re-weight the parameters. Readers can see what was done from the WinBUGS code provided in *Appendix 5*, which is annotated to make these adjustments clear.

Prior distributions

Vague normal priors were placed on the log incidence rate in the LaMontagne study[192] in age group 1 and GP setting: $\ln(\lambda_{1,GP,1}) \sim N(0,100^2)$, and also on the rate ratios ρ_s for setting s relative to the GP setting, and γ_a for age group a relative to the 16- to 17-year-old group, and for the ratios η_s in re-infection rate to infection rate in setting s: ρ_a, γ_s, $\eta_s \sim N(0,100^2)$.

Priors for the duration of infection and proportion symptomatic were as follows: proportion symptomatic $\varphi \sim Beta(1,1)$; $\Delta_A \sim n(0,100^2)$, $\Delta_S \sim$ Uniform(0.0767,0.1533), i.e. uniform between 4 and 8 weeks. Information on the proportion of patients at recruitment in the GP setting, $p_{a,GP}$ in the LaMontagne study[192] who were in the re-infection group (re-weighted to account for disproportionate inclusion into the study of initially CT+ women) was introduced via informative Beta priors, derived from table 1 in LaMontagne *et al.*[192] For example, in women aged 16–19 years, who were tested at GP clinics, 663 + 137 = 800 were CT– and 45 + 48 = 93 were CT+. So the correct weights for the infection and re-infection groups are 800/893 and 93/893, respectively. We repeated the same calculation for women aged 20–24 years, and

we assume that the weights are constant within these two age groups. Although testing and treatment every 6 months interferes with the natural history, CT+ women are subsequently placed in the re-infection group so this does not bias the results.

Information on two parameters, Δ_S, and the proportion of women in the re-infections group, $p_{a,GP}$, was introduced via informative priors rather than through the data likelihood. This prevents these 'data' from contributing directly to the global goodness of fit assessment. The decision to treat these inputs differently was because the source of evidence on the first was expert clinical knowledge that was quite unrelated to the other sources of data in the synthesis, whereas the second was local to the LaMontagne study.[192] We were therefore interested less in the 'goodness of fit' of this information, and more in the goodness of fit of the other data, conditional on the priors we assigned to these parameters. In addition, we applied the 'Cut Function' to both these parameters (see *Chapter 3*), so in these cases the posterior for the parameter is the same as the informative prior.

Data likelihoods

The age-specific prevalence data $D_{a,pop}$ in *Table 10* was given a normal likelihood on the logit scale: $\text{logit}(D_{a,pop}) \sim N(\pi_{a,pop}, V_{a,pop})$, with the variance calculated from the 95% CIs. The setting-specific ORs in *Table 11* were handled in the same way: $\text{logit}(OR_s) \sim N(\gamma_s, V_s)$. The data on duration of asymptomatic infection (see *Table 8*) was entered as a normal likelihood: $Dur_A \sim N(\Delta_A, V_A)$.

The numbers infected in *Table 9* denoted 'r' are considered as having a binomial likelihood, with parameters $\kappa(t)_{a,s,i}$, and denominators also shown in the table so that $r_{a,s,i} \sim Bin(\kappa_{a,s,i}(0.5), n_{a,s,i})$. The number of symptomatic infections, ($r = 26$), reported by Geisler *et al.*[73] is binomially distributed with parameter φ and denominator 115.

The WinBUGS code in *Appendix 5* consists of the priors and likelihoods as described above, and the functional relationships described exactly as in *Figure 8* and in the text.

Appendix 5 WinBUGS code and data for the *Chapter 5* analysis

```
model {

# Duration
# Duration of asymptomatic infection
dura ~ dnorm(0,0.0001) I(0,)
D ~ dnorm(dura,pr.D)
dev.D <- pr.D*pow((dura-D),2)  # Duration Deviance
lambdaC <- 1/dura              # clearance rate

# Duration of symptomatic infection
durs ~ dunif(0.0767,0.1533)  # symptomatic, 4-8 weeks
#durs ~ dunif(0.0577,0.2308) #symptomatic, 3-12 weeks sensitivity
# analysis
cut.durs <- cut(durs)        # pevents updating of durs
lambdaS <- 1 / cut.durs      # clearance rate

# Probability ct is symptomatic
r.phi ~ dbin(phi,n.phi)  phi ~ dbeta(1,1)
dev.phi <- 2 * (r.phi * log(r.phi / (phi *  n.phi)) +  (n.phi -
r.phi) *
          log((n.phi - r.phi) / (n.phi - (n.phi * phi))))  #
# Deviance

# Mean duration of infection
dur <- phi*cut.durs + (1-phi)*dura



# Incidence model
# Likelihood and model
for (i in 1:2) { # loop over rate, 1=infection, 2=reinfection
 for (s in 1:3) { # loop over setting - GP=1, fP=2, GUM=3
  for (a in 1:3) { # loop over age group:1=16,17;2=18-20, 3=21-24
   lambda[i,s,a] <- equals(i,1)*(gamma[a] * rho[s] * lambda111) +
                    equals(i,2)*(gamma[a] * rho[s] * lambda111 *
eta[s])
   r.inc[i,s,a] ~ dbin(theta.inc[i,s,a],n.inc[i,s,a])
   theta.inc[i,s,a] <- phi * (
                       (1 - (lambdaS + lambda[i,s,a] *
                        exp(-(lambdaS + lambda[i,s,a]) * 0.5)) /
                       (lambdaS + lambda[i,s,a]))
                      ) + (1 - phi) * (
                       (1 - (lambdaC + lambda[i,s,a] *
                      exp(-(lambdaC + lambda[i,s,a]) * 0.5)) /
                       (lambdaC + lambda[i,s,a]))
                                        )
  dev.inc.rat[i,s,a] <- 2 * (r.inc[i,s,a] * log(r.inc[i,s,a] /
                        (theta.inc[i,s,a] * n.inc[i,s,a])) +
(n.inc[i,s,a]-
                        r.inc[i,s,a]) * log((n.inc[i,s,a] -
r.inc[i,s,a]) /
                       (n.inc[i,s,a] - (n.inc[i,s,a] *
theta.inc[i,s,a]))))
   }
  }
 }
sumdev.inc.rat <- sum(dev.inc.rat[ , , ])   # Deviance, # LaMontagne

#priors for log incidence parameters
```

```
loglambda111 ~ dnorm(0,.0001)                    #Age 16-17, GP,
# infection
log(lambda111) <- loglambda111

for (s in 1:3) {
 logeta[s] ~ dnorm(0,.0001)     # GP=1, fP=2, GUM=3
 log(eta[s]) <- logeta[s]
 }
logrho[1] <- 0                  # GPtheta.inc
for (s in 2:4) {
 logrho[s] ~ dnorm(0,.0001) # 2=fp, 3=GUM, 4=pop, rel to GP
 }
for (s in 1:4) {
 log(rho[s]) <- logrho[s]
 }

loggamma[1] <- 0            #1 = 16-17
for (a in 2:5) {
 loggamma[a] ~ dnorm(0,.01) #2=18-19, 3=20-24,4=25-29,5=30-44 rel to
# 16-17
}
for (a in 1:5) {
 log(gamma[a]) <- loggamma[a]
 }

# setting specific Odds ratios from Adams
for (s in 1:3) {                            #s+1= 2=FP, 3=GUM, 4=pop,
# rel to GP
   r.OR[s] ~ dnorm(logrho[s+1],pr.OR[s])
   dev.OR[s] <- pr.OR[s] * pow((r.OR[s]-logrho[s+1]),2)  # OR
# deviance
}
sumdev.OR <- sum(dev.OR[])

# overall age-specific incidence in population, 1=16-17, 2=18-19,
# 3=20-24,
#                                              4=25-29, 5=30-44
for (a in 1:3) {
 p[a] ~ dbeta(aprior[a],bprior[a])
 cut.p[a] <- cut(p[a])                      # prevents updating p[a]
 lamda.F[a] <- rho[4]*((1-cut.p[a])*lambda[1,1,a] +
               cut.p[a]*lambda[2,1,a])
 lamda.pop[a] <- lamda.F[a] / (1 - lamda.F[a] * dur)
 }
for (a in 4:5) {
 lamda.pop[a] <- lamda.pop[1] * gamma[a]
 }

# Age-specific Prevalence from Adams, 1=18-19, 2=20-24, 3=25-29,
# 4=30-44
for (a in 1:4) {
 Lprev[a]~dnorm(lprev[a+1],pr.Lprev[a])
 dev.prev[a] <- pr.Lprev[a] * pow((Lprev[a]-lprev[a+1]),2)
 }
for (a in 1:5) {
 lprev[a] <- logit(prev[a])
 }

for (a in 1:3) {
 prev[a] <- min(.999,(dur * lamda.pop2[a]))
}
```

```
prev[4] <- prev[3] * gamma[4]/gamma[3]
prev[5] <- prev[3] * gamma[5]/gamma[3]

sumdev.prev <- sum(dev.prev[])

sumdev <- dev.D + sumdev.inc.rat + sumdev.OR + sumdev.prev + dev.phi


# incidence in difference age bands, using population weights N[]
for (a in 16:17) {inc[a] <- lamda.pop[1]}
for (a in 18:20) {inc[a] <- lamda.pop[2]}
for (a in 21:24) {inc[a] <- lamda.pop[3]}
for (a in 25:29) {inc[a] <- lamda.pop[4]}
for (a in 30:44) {inc[a] <- lamda.pop[5]}

lamda.pop2[1] <- lamda.pop[1]  # age 16,17
lamda.pop2[2] <- lamda.pop[2]  # age 18,19.
lamda.pop2[3] <- inprod(inc[20:24],N[20:24])/sum(N[20:24])  # age 20-
# 24

inc1624 <- inprod(inc[16:24],N[16:24])
inc2544 <- inprod(inc[25:44],N[25:44])
inc1644 <- inprod(inc[16:44],N[16:44])

lamda.pop2[4] <- inc1624 / sum(N[16:24]) # age 16-24
lamda.pop2[5] <- inc2544 / sum(N[25:44]) # age 25-44
lamda.pop2[6] <- inc1644 / sum(N[16:44]) # age 16-44

# other population summaries
prop.treat.1624.2002 <- 31510 / inc1624
prop.treat.1624.2003 <- 34660 / inc1624
prop.asymp.clin.1624.2003 <- 1 - (phi / prop.treat.1624.2003)

# prevalence in reconstructed age groups
for (a in 1:6) {
 prev.pop2[a] <- dur * lamda.pop2[a]
 }
} # end of program


Data
list(

# duration
D=1.36, pr.D=59.17,    # from Duration paper 2-Class estimate

# incidence
r.inc=structure(.Data=c(
4,3,4, 9,5,7, 5,16,9,
5,7,10, 13,12,5, 6,15,5
),.Dim=c(2,3,3)),

n.inc=structure(.Data=c(
73,195,188, 194,273,201, 102,235,245,
14,65,79, 95,127,63, 40,139,81
),.Dim=c(2,3,3)),

# priors for infection/reinfection weights, 16-17, 18-20, 21-24
aprior=c(46.5,73.1,106.4), bprior=c(400,634.6,938.4),

# Logit Prevalence
```

```
Lprev=c(-2.987,-3.41,-4.185,-4.82),
pr.Lprev=c(19.3,20.8,18.4,17.2),

# Setting-specific Log Odds Ratios. Fp,GUM, pop
r.OR= c(0.239,0.871, -.511),
pr.OR= c(67.03,35.21, 17.28),

# symptomatic ct - Geisler
r.phi = 26, n.phi = 115,

# Population sizes from census, age =1...44 - 2002
N=c(NA,NA,NA,NA,NA, NA,NA,NA,NA,NA, NA,NA,NA,NA,NA,
305500,306300,296400,291400,294800,
310100,313900,305600,294700,295000,
304100,317000,329600,349600,370300,
380900,376900,387800,390900,399400,
401200,402600,398700,391900,381900, 370900,356200,349000,343800)
)

# Initial values - 1
# Duration
list(dura=1,  durs=.115,

# incidence
loglambda111=-2.5,
logeta=c(1,.5,.5),
logrho=c(NA, 1, 1.5,-.5),
loggamma=c(NA,-.2,-.5,-.7,-.8),

# probability symptomatic
phi = 0.5,

# proportion re-infection
p=c(.1, .1, .1)
)


# Initial values - 2
# Duration
list(dura=.2,  durs=.112,

# incidence
loglambda111=-4,
logeta=c(.5,1,1),
logrho=c(NA, 0, 0,0),
loggamma=c(NA,0,0,0,0),

# probability symptomatic
phi = .1,

# proportion re-infection
p=c(.3, .05, .2)
)
```

Appendix 6 Sensitivity analyses in *Chapter 6*

We performed three sets of sensitivity analysis.

First, the proportion of patients in the Scholes *et al.* trial[13] who were assumed to have been tested during the trial was lowered from 32% to 15%. This had no measurable effect on results of the final synthesis (results not shown).

The effect on key model results of changing the infection and reinfection rates for both the one- and two-rate models is shown in *Table 46*. The results show clearly that in the full synthesis the assumed infection/re-infection rates have a minimal impact on the results altering the estimates of κ, the probability that an incident CT episode causes PID, by at most a multiplicative factor of 4%.

In summary, we found that results from the full synthesis models were insensitive to these modelling assumptions.

TABLE 46 Effect on key results of altering the assumed infection and reinfection rates, *Chapter 6*

Infection rate (Scholes 1996,[13] Ostergaard 2000[12])	Re-infection rate (POPI 2010,[18] Rees 1980[199])	Mean residual deviance	Causal rate of PID (95% CrI)	Probability CT causes clinical PID (95% CrI)	Proportion prevented by screening (95% CrI)
One-rate model: all controlled studies					
0	0	10.8	0.19 (0.06 to 0.36)	0.16 (0.06 to 0.26)	0.61 (0.55 to 0.67)
0	0.05	10.7	0.19 (0.06 to 0.36)	0.16 (0.06 to 0.26)	0.61 (0.55 to 0.67)
0	0.1	10.6	0.19 (0.06 to 0.35)	0.16 (0.06 to 0.26)	0.61 (0.55 to 0.67)
0.05	0.1	10.6	0.19 (0.06 to 0.35)	0.16 (0.06 to 0.26)	0.61 (0.55 to 0.67)
0.05	0.15	10.7	0.19 (0.06 to 0.34)	0.16 (0.06 to 0.25)	0.61 (0.55 to 0.68)
0.05	0.20	10.8	0.19 (0.06 to 0.34)	0.15 (0.06 to 0.25)	0.61 (0.55 to 0.68)
0.1	0.15	10.6	0.18 (0.06 to 0.31)	0.15 (0.05 to 0.24)	0.61 (0.56 to 0.68)
0.1	0.20	10.7	0.18 (0.05 to 0.31)	0.15 (0.05 to 0.24)	0.62 (0.56 to 0.68)
Two-rate models: all controlled studies – 60 days					
0	0	10.3	δ_1 0.74 (0.07 to 1.72) δ_2 0.16 (0.03 to 0.33)	0.20 (0.09 to 0.31)	0.40 (0.13 to 0.68)
0	0.05	10.4	δ_1 0.62 (0.05 to 1.46) δ_2 0.17 (0.03 to 0.34)	0.19 (0.09 to 0.30)	0.43 (0.16 to 0.69)
0	0.1	10.8	δ_1 0.48 (0.03 to 1.18) δ_2 0.17 (0.04 to 0.34)	0.18 (0.08 to 0.28)	0.48 (0.20 to 0.70)

continued

TABLE 46 Effect on key results of altering the assumed infection and reinfection rates, *Chapter 6* (*continued*)

Infection rate (Scholes 1996,[13] Ostergaard 2000[12])	Re-infection rate (POPI 2010,[18] Rees 1980[199])	Mean residual deviance	Causal rate of PID (95% CrI)	Probability CT causes clinical PID (95% CrI)	Proportion prevented by screening (95% CrI)
0.05	0.1	10.6	δ_1 0.32 (0.02 to 0.68) δ_2 0.18 (0.04 to 0.35)	0.16 (0.07 to 0.26)	0.54 (0.30 to 0.72)
0.05	0.15	10.9	δ_1 0.29 (0.02 to 0.64) δ_2 0.18 (0.05 to 0.35)	0.16 (0.07 to 0.26)	0.55 (0.32 to 0.72)
0.05	0.20	11.2	δ_1 0.25 (0.01 to 0.60) δ_2 0.18 (0.05 to 0.35)	0.16 (0.07 to 0.25)	0.57 (0.34 to 0.72)
0.1	0.15	11.0	δ_1 0.18 (0.01 to 0.38) δ_2 0.19 (0.05 to 0.36)	0.15 (0.06 to 0.25)	0.61 (0.43 to 0.73)
0.1	0.20	11.1	δ_1 0.17 (0.01 to 0.37) δ_2 0.19 (0.05 to 0.36)	0.15 (0.06 to 0.24)	0.61 (0.44 to 0.73)

Appendix 7 Development of functional relationships (see *Chapter 6*)

Note that the notation used in the chapter is extended here to facilitate a full mathematical description of the *non-homogeneous* model. Here we develop the mathematical relationships between the model parameters and data estimands.

Note: in our analyses the values of λ^{l}, u', u, θ^{CT-} and hence any functions of them such as $p_{ij}(t)$ are all study dependent but the s subscript has been dropped for ease of reading.

Expression in a homogeneous model

In a single rate model θ^{CT+} is constant over time and the data from women who are CT+ and the women who are CT– at baseline in study s inform the transition probabilities $p_{13}(t)$ and $p_{23}(t)$, respectively, where t is the mean follow-up time of the study. These transition probabilities are obtained by solving equation 19 numerically in each iteration of the MCMC simulation using the Runge–Kutta method in WBDiff.

Expression in a two-step piecewise homogeneous model

We use a 1-day discrete time approximation to the underlying continuous process to simplify the calculations. It is necessary to extend the notation used in the paper. Let $\pi_{i,h}$ be the state occupancy proportion for state i in day h of the study. The proportion in state $i=1$ is further subdivided into the proportion $\pi^{hi}_{i,h}$, who are subject to developing PID caused by CT at rate δ_1, and the proportion $\pi^{low}_{i,h}$ who develop PID caused by CT at rate δ_2 where $\pi_{1,h}=\pi^{low}_{1,h}+\pi^{hi}_{1,h}$.

Finally, p^{hi}_{ij} and p^{low}_{ij} are the daily transition probabilities from state i to state j, for women who are subject to developing PID caused by CT at rates δ_1 and δ_2, respectively. p^{hi}_{ij} and p^{low}_{ij} are calculated using equation 19, and solved using WBDiff, with the appropriate expression for θ^{CT+}_{s} substituted into equation 18. The observed data inform the proportions $\pi_{3,H}$ calculated for each arm where H is equal to the follow-up period of the study in days. So $\pi_{3,H}$ is the proportion of patients in state 3 at the end of the study, which is the parameter in the binomial likelihood. Equations for the state occupancy proportions on day h of the study can be written as follows:

$$\begin{aligned}
\pi_{1,h} &= \pi^{low}_{1,h-1}\cdot p^{low}_{11} + \pi^{hi}_{1,h-1}\cdot p^{hi}_{11} + \pi_{2,h-1}\cdot p^{hi}_{21} \\
\pi_{2,h} &= \pi_{2,h-1}\cdot p^{hi}_{22} + \pi^{low}_{1,h-1}\cdot p^{low}_{12} + \pi^{hi}_{1,h-1}\cdot p^{hi}_{12} \\
\pi_{3,h} &= 1-\pi_{2,h}-\pi_{1,h}
\end{aligned} \tag{74}$$

where $\pi_{1,0}=\pi^{low}_{1,0}+\pi^{hi}_{1,0}$, $\pi_{2,0}=1-\pi_{1,0}$, $\pi_{3,0}=0$.

It just remains to specify equations for $\pi^{hi}_{1,h}$ and $\pi^{low}_{1,h}$, which depend on whether the study is clinic or screening based. To speed up the processing time, the code is written in WBDev.[160]

Clinic-based studies

$$\left.\begin{aligned}\pi^{hi}_{1,h} &= \pi_{1,h} \\ \pi^{hi}_{1,h} &= \sum_{i=(h+1-B)}^{h} (\pi_{2,i-1}\cdot p^{hi}_{21}) p^{hi^{(h-i)}}_{11} \\ \pi^{low}_{1,h} &= \pi_{1,h} - \pi^{hi}_{2,h}\end{aligned}\right\} \begin{aligned} h &= 1,\ldots, B \\ h &= B+1,\ldots, H\end{aligned} \tag{75}$$

Recall from the main text that B is the number of days for which patients are subject to the rate δ_1 of PID caused by CT after entering state 1 if they do not leave during this period. So, in clinic-based studies, all women in state 1 during the first B days are subject to the rate δ_1. Equation 19 shows the proportion of women who are in state 1 and progress at rate δ_1 for each day h in the study after B days have elapsed. Equation 18 sums over all patients who have moved to state 1 in the last B days multiplied by the probability they have remained there until day h. In the case arm, $\pi^{hi}_{1,0} = 1$ and $\pi^{low}_{1,0}$ and $\pi_{2,0}$ equal 0. In the control arm, $\pi^{hi}_{1,0}$ and $\pi^{low}_{1,0}$ equal 0, and $\pi_{2,0}$ equals 1.

Screening-based studies

In screened populations women who are CT+ have already been infected for a period of time before recruitment. We assume that a proportion, ϕ, of incident CT cases are symptomatic and treated, clearing at rate λ^T, and the remaining infections, $1-\phi$, are asymptomatic and clear at rate λ^C. The probability w_b that a recruited patient was recruited exactly b days after infection is:

$$\begin{aligned}\omega_b = {} & \varphi\cdot\left(\exp\left(-\lambda^T\cdot\frac{(b-1)}{365}\right)-\exp\left(-\lambda^T\cdot\frac{b}{365}\right)\right) \\ & +(1-\varphi)\cdot\left(\exp\left(-\lambda^C\cdot\frac{(b-1)}{365}\right)-\exp\left(-\lambda^C\cdot\frac{b}{365}\right)\right)\end{aligned} \tag{76}$$

The proportions in the high and low rate groups are:

$$\pi^{hi}_{1,h} = \sum_{i=h}^{B} \omega_{B+1-i}\pi_{1,0}\cdot p^{h}_{11} + \sum_{i=1}^{h} (\pi_{2,i-1}\cdot p^{hi}_{21}) p^{hi^{(h-i)}}_{11} \quad \Big\} h = 1,\ldots, B \tag{77}$$

$$\left.\begin{aligned}\pi^{hi}_{1,h} &= \sum_{i=(h+1-B)}^{h} (\pi_{2,i-1}\cdot p^{hi}_{21}) p^{hi^{(h-i)}}_{11} \\ \pi^{low}_{1,h} &= \pi_{1,h} - \pi^{hi}_{2,h}\end{aligned}\right\} h = B+1,\ldots, H \tag{78}$$

Equation 77 calculates the proportion of women who are in state 1 and subject to the first causal progression rate δ_1 on each day h of the study from 1 to B. The first term is the proportion of women who began the study as CT+ and were infected within 60 days of day h, multiplied by the probability they have remained in state 1 until day h. This term is not included in equation 78 because, after B days, all of the remaining patients move to the δ_2 rate. The second term sums over all patients who have moved to state 1 by day h, multiplied by the probability they have remained there until day h. Equation 78 is the same as for clinic based studies. In the CT+ group, $\pi^{hi}_{1,0} = \sum_{b=1}^{B} \omega_b$, $\pi^{low}_{1,0} = 1-\pi^{hi}_{1,0}$ and $\pi_{2,0}$ equals 0. In the CT– group, $\pi^{hi}_{1,0}$ and $\pi^{low}_{1,0}$ equal 0, and $\pi_{2,0}$ equals 1.

Proportion of *Chlamydia trachomatis*-related pelvic inflammatory disease that is prevented by screening

We consider an annual screening programme, in which women are screened at exactly 365-day intervals. We derive an expression for the proportion of episodes of PID prevented by annual screening in women who become infected with CT. We assume that treatment is administered 14 days after testing (the results are fairly insensitive to different reasonable assumptions about this time period). Women are screened within a year of infection, with each day having an equal probability (1/365) of this occurring. The expression differs between the one- and two-rate models.

One-rate model

The probability that an episode of PID is caused by a CT episode exactly b days after infection is:

$$\begin{aligned}\kappa_b = \varphi.\left(\left(1-\exp\left(-\left(\lambda^T+\delta\right).t_{u_0u_b}\right)\right)-\left(1-\exp\left(-\left(\lambda^T+\delta\right).t_{u_0u_{(b-1)}}\right)\right).\frac{\delta}{\lambda^T+\delta}\right.\\ \left.+\left(1-\varphi\right).\left(\left(1-\exp\left(-\left(\lambda^C+\delta\right).t_{u_0u_b}\right)\right)-\left(1-\exp\left(-\left(\lambda^C+\delta\right).t_{u_0u_{(b-1)}}\right)\right).\frac{\delta}{\lambda^C+\delta}\right.\end{aligned} \tag{79}$$

Under the assumptions outlined above, the probability κ^{SCN} a CT infection causes an episode of PID in a woman who is screened at random annually is:

$$k^{SCN}=\sum_{b=1}^{14}\kappa_{b'}+\sum_{b=15}^{379}\kappa_{b'}\left(1-\frac{(b-14)}{365}\right) \tag{80}$$

And the proportion of PID episodes that are caused by CT, in women with CT, which are prevented by annual testing, χ, equals:

$$\chi=1-\frac{\kappa^{SCN}}{\kappa} \tag{81}$$

Two-rate model

In a two-rate model the probability that an episode of PID develops exactly b days after infection equals:

$$\kappa_b=\left.\begin{array}{l}\varphi.\left(\exp\left(-\left(\lambda^T+\delta_1\right).t_{u_0u_{(b-1)}}-\exp\left(-\left(\lambda^T+\delta_1\right).t_{u_0u_b}\right)\right).\frac{\delta_1}{\lambda^T+\delta_1}+\\ (1-\varphi).\left(\exp\left(-\left(\lambda^C+\delta_1\right).t_{u_0u_{(b-1)}}-\exp\left(-\left(\lambda^C+\delta_1\right).t_{u_0u_b}\right)\right).\frac{\delta_1}{\lambda^C+\delta_1}\end{array}\right\}\ b\le B$$

$$\left.\begin{array}{l}\varphi.\left(\exp\left(-\left(\lambda^T+\delta_2\right).t_{u_1u_{(b-1)}}-\exp\left(-\left(\lambda^T+\delta_2\right).t_{u_1u_b}\right)\right).\frac{\delta_2}{\lambda^T+\delta_2}+\\ (1-\varphi).\left(\exp\left(-\left(\lambda^C+\delta_2\right).t_{u_1u_{(b-1)}}-\exp\left(-\left(\lambda^C+\delta_2\right).t_{u_1u_b}\right)\right).\frac{\delta_2}{\lambda^C+\delta_2}\end{array}\right\}\ b>B \tag{82}$$

Where B = 30, 60, 90 days, the proportion of PID cases prevented by screening is calculated from equation 80 and 81 as before.

Appendix 8 WinBUGS code for the two-rate model (see *Chapter 6*)

```
model {
# chlamydia informative priors
r.phi ~ dbin(phi,129)
lambda.c ~ dnorm(0.743,193)
dur.symp ~ dunif(0.077,0.15)  # 4 - 8 weeks
lambda.t <- 1 / cut(dur.symp)

# Prospective PID analysis constant progression rate
# Likelihood
for (s in 1:4) {
 for (i in 1:2) {
  r[s,i] ~ dbin(p[s,i],n[s,i])
  }
 }
# transistion probability calulations
for (s in 1:4) {
 for (i in 1:2) {
  for(j in 1:3) {
   theta[s,1,index[i,j]] <- lambda[s,1,i,j]
   theta[s,2,index[i,j]] <- lambda[s,2,i,j]
   }
  }
 lambda[s,1,1,1] <- - lambda[s,1,1,2] - lambda[s,1,1,3]
 lambda[s,1,1,3] <- theta.CTpos1[s]
 lambda[s,1,2,2] <- - lambda[s,1,2,1] - lambda[s,1,2,3]
 lambda[s,1,2,3] <- theta.CTneg[s]

 lambda[s,2,1,1] <- - lambda[s,2,1,2] - lambda[s,2,1,3]
 lambda[s,2,1,3] <- theta.CTpos2[s]
 lambda[s,2,2,2] <- - lambda[s,2,2,1] - lambda[s,2,2,3]
 lambda[s,2,2,3] <- theta.CTneg[s]
 }

# Infection rate
for (s in 1:2) {
 lambda[s,1,2,1] <- 0.0
 lambda[s,2,2,1] <- 0.0
 }

for (s in 3:4) {
 lambda[s,1,2,1] <- 0.0
 lambda[s,2,2,1] <- 0.0
 }

# Clearance + treatment rate
lambda[1,1,1,2] <- lambda.c + 0.64
lambda[1,2,1,2] <- lambda.c + 0.64

lambda[2,1,1,2] <- lambda.c + 0.43
lambda[2,2,1,2] <- lambda.c + 0.43

lambda[3,1,1,2] <- lambda.c + 0.32
lambda[3,2,1,2] <- lambda.c + 0.32

lambda[4,1,1,2] <- lambda.c + 0.32
lambda[4,2,1,2] <- lambda.c + 0.32

# Models for rates
for (s in 1:4) {
 theta.CTpos1[s] <- alpha[s] + delta[1]
```

```
 theta.CTpos2[s] <- alpha[s] + delta[2]
 theta.CTneg[s] <- alpha[s]
 }

# Other data
r.sch ~ dbin(pi.sch,n.sch) # Scholes - ct prevalence
r.ost ~ dbin(pi.ost,n.ost) # Ostergaard - ct prevalence

# Priors
delta[1] ~ dexp(0.00001)
delta[2] ~ dexp(0.00001)
for (s in 1:4) {
 alpha[s] ~ dexp(0.00001)
 }
pi.sch ~ dbeta(1,1)
pi.ost ~ dbeta(1,1)
phi ~ dbeta(1,1)

# Calculation of parameters in the likekihood
for (s in 1:4) {
 solution1d[s,1,1:dim] <-
three.state(init[1:dim],time,theta[s,1,1:n.par],
                    origin, tol)
 solution1d[s,2,1:dim] <-
three.state(init[1:dim],time,theta[s,2,1:n.par],
                    origin, tol)
 for (i in 1:2) {
  for (z in 1:6) {
   vectorforwbdev[s,i,z] <- solution1d[s,1,z]
   vectorforwbdev[s,i,z+6] <- solution1d[s,2,z]
   }
  vectorforwbdev[s,i,13] <- round(t[s] * 365)
  vectorforwbdev[s,i,14] <- B
  vectorforwbdev[s,i,15] <- lambda.c
  vectorforwbdev[s,i,16] <- lambda.t
  vectorforwbdev[s,i,17] <- phi
  vectorforwbdev[s,i,18] <- clinorscreen[s]
  vectorforwbdev[s,i,19] <- caseprop[s,i]

  p[s,i] <- generatep(vectorforwbdev[s,i,1:19])
  }
 caseprop[s,2] <- 0
 }
caseprop[1,1] <- 1
caseprop[2,1] <- 1
caseprop[3,1] <- pi.sch
caseprop[4,1] <- pi.ost

clinorscreen[1] <- 0 # clinic based study
for (s in 2:4) {
 clinorscreen[s] <- 1   # screening studies
 }

# Residual Deviance
for (s in 1:4) {
 for (i in 1:2) {
  dev[s,i] <- 2 * (r[s,i] * log(r[s,i] / (p[s,i] * n[s,i])) +
                  (n[s,i] - r[s,i]) * log((n[s,i] - r[s,i]) /
                  (n[s,i] - (n[s,i] * p[s,i]))))
  }
 }
```

```
dev.sch <- 2 * (r.sch * log(r.sch / (pi.sch * n.sch)) +
                (n.sch - r.sch) * log((n.sch - r.sch) /
                (n.sch - (n.sch * pi.sch))))

dev.ost <- 2 * (r.ost * log(r.ost / (pi.ost * n.ost)) +
                (n.ost - r.ost) * log((n.ost - r.ost) /
                (n.ost - (n.ost * pi.ost))))

sumdev <- sum(dev[ , ]) + dev.sch + dev.ost


# Results
kappa  <- (1 - phi) * (
          (1 - (exp( - (lambda.c + delta[1]) * (B / 365)))) *
           delta[1] / (lambda.c + delta[1]) +
           exp( - (lambda.c + delta[1]) * (B / 365)) *
           delta[2] / (lambda.c + delta[2])
           ) +
           phi * (
          (1 - (exp( - (lambda.t + delta[1]) * (B / 365)))) *
           delta[1] / (lambda.t + delta[1]) +
           exp( - (lambda.t + delta[1]) * (B / 365)) *
           delta[2] / (lambda.t + delta[2])
           )

# proportion of PIDs prevented by annual screening
# assumes two-week delay between test and treatment
for (i in 1:B) {
 temp1[i] <- phi * (
             ((1 - exp(-(lambda.t + delta[1]) * (i/365))) -
              (1 - exp(-(lambda.t + delta[1]) * ((i-1)/365)))) *
              (delta[1] / (lambda.t + delta[1])) *
              (1 - (max(0,(i-14)) / 365))) +
             (1 - phi) * (
             ((1 - exp(-(lambda.c + delta[1]) * (i/365))) -
              (1 - exp(-(lambda.c + delta[1]) * ((i-1)/365)))) *
              (delta[1] / (lambda.c + delta[1])) *
              (1 - (max(0,(i-14)) / 365)))
 }
for (i in B+1:379) {
 temp1[i] <- phi * (
             ((1 - exp(-(lambda.t + delta[2]) * (i/365))) -
              (1 - exp(-(lambda.t + delta[2]) * ((i-1)/365)))) *
              (delta[2] / (lambda.t + delta[2])) *
              (1 - (max(0,(i-14)) / 365))) +
             (1 - phi) * (
             ((1 - exp(-(lambda.c + delta[2]) * (i/365))) -
              (1 - exp(-(lambda.c + delta[2]) * ((i-1)/365)))) *
              (delta[2] / (lambda.c + delta[2])) *
              (1 - (max(0,(i-14)) / 365)))
 }
prop.prevent <- 1 - (sum(temp1[ ]) / kappa)

# Bayesian p-value
test <- delta[1] - delta[2]
B.p <- step(test)
}

# Data
list(
```

```
# PID (1 month) prospective
# 1. Rees
# 2. POPI
# 3. Scholes
# 4. Ostergaard

r=structure(.Data=c(
8,3,
7,1,
33,7,
20,9
),.Dim=c(4,2)),

n=structure(.Data=c(
67,62,
74,63,
1598,645,
487,443
),.Dim=c(4,2)),
t = c(0.125,1,1,1),
B = 60,   # If this were set above 365 WBDEV code would need changing
r.sch = 44,
n.sch = 645,
r.ost = 43,
n.ost = 867,
r.phi=30,
time = 0.00274,

# forward equations
dim=6,origin=0,tol=1.0E-4, init=c(1,0,0, 0,1,0),n.par=6,
index=structure(.Data=c(1,2,3,
                        4,5,6), .Dim=c(2,3))
)


# Initial values - 1
list(
# Prospective
phi = 0.23, lambda.c = 0.74, delta = c(0.2,0.1),
alpha = c(0.01,0.01,0.01,0.01),
pi.sch = 0.051, pi.ost = 0.07, dur.symp = 0.1
)

# Initial values - 2
list(
# Prospective
phi = 0.7, lambda.c = 2, delta = c(0.1,0.6),
alpha = c(0.15,0.15,0.15,0.15),
pi.sch = 0.2, pi.ost = 0.2, dur.symp = 0.145
)
```

WBDiff Program to calculate the transition probabilities

```
MODULE WBDiffThreeState;

 IMPORT
  WBDiffODEMath,
  Math;

 TYPE
  Equations = POINTER TO RECORD (WBDiffODEMath.Equations) END;
```

```
  Factory = POINTER TO RECORD (WBDiffODEMath.Factory) END;

 CONST

  nEq = 6;
  nSt = 4; (* one higher as arrays start at zero*)


 VAR
  fact-: WBDiffODEMath.Factory;

 PROCEDURE (e: Equations) Derivatives (IN theta, C: ARRAY OF REAL; n: INTEGER; t:
REAL;
                                       OUT dCdt: ARRAY OF REAL);
 VAR

  index: ARRAY nSt, nSt OF INTEGER;

 BEGIN

(* define index of parameters (look-up table) *)
  index[1,1] := 0;
  index[1,2] := 1;
  index[1,3] := 2;
  index[2,1] := 3;
  index[2,2] := 4;
  index[2,3] := 5;

(* define system of nEq Differential Equations *)
  dCdt[index[1,1]]:= C[index[1,1]]*theta[index[1,1]] + C[index[1,2]]*theta[index[2,1]];
  dCdt[index[1,2]]:= C[index[1,1]]*theta[index[1,2]] + C[index[1,2]]*theta[index[2,2]];
  dCdt[index[1,3]]:= C[index[1,1]]*theta[index[1,3]] + C[index[1,2]]*theta[index[2,3]];

  dCdt[index[2,1]]:= C[index[2,1]]*theta[index[1,1]] + C[index[2,2]]*theta[index[2,1]];
  dCdt[index[2,2]]:= C[index[2,1]]*theta[index[1,2]] + C[index[2,2]]*theta[index[2,2]];
  dCdt[index[2,3]]:= C[index[2,1]]*theta[index[1,3]] + C[index[2,2]]*theta[index[2,3]];


  END Derivatives;

  PROCEDURE (equations: Equations) SecondDerivatives (IN theta, x: ARRAY OF REAL;
                                  numEq: INTEGER; t: REAL; OUT d2xdt2: ARRAY OF REAL);
  BEGIN
   HALT(126)
  END SecondDerivatives;

PROCEDURE (equations: Equations) Jacobian (IN theta, x: ARRAY OF REAL;
                    numEq: INTEGER; t: REAL; OUT jacob: ARRAY OF ARRAY OF REAL);
  BEGIN
   HALT(126)
  END Jacobian;

PROCEDURE (f: Factory) New (option: INTEGER): WBDiffODEMath.GraphNode;
  VAR
   equations: Equations;
   node: WBDiffODEMath.GraphNode;
  BEGIN
   NEW(equations);
   node := WBDiffODEMath.New(equations, nEq);
   RETURN node
```

```
 END New;

 PROCEDURE Install*;
 BEGIN
  WBDiffODEMath.Install(fact)
 END Install;

 PROCEDURE Init;
 VAR
  f: Factory;
 BEGIN
  NEW(f); fact := f
 END Init;

BEGIN
 Init
END WBDiffThreeState.
```

WBDEV code to calcuate the Binomial likelihood probabilities of PID for the 2-rate model

```
MODULE WBDevgeneratep;

 IMPORT
  WBDevScalar,
  Math;

 TYPE
   Function = POINTER TO RECORD (WBDevScalar.Node) END;
   Factory = POINTER TO RECORD (WBDevScalar.Factory) END;

 VAR
   fact-: WBDevScalar.Factory;

PROCEDURE (func: Function) DeclareArgTypes (OUT args: ARRAY OF CHAR);
 BEGIN
  args := "v";
 END DeclareArgTypes;

PROCEDURE calculation (func: Function; OUT output: REAL);
 VAR
  term1, term2, omega, omega_cum,pi_hi, pi_low:                ARRAY 366 OF REAL;
  pi:                                                          ARRAY 4,366+1 OF REAL;

  p_hi,  p_low:                                                ARRAY 3,4 OF REAL;
  H,B,h,i,b:                                                   INTEGER;
  Hin, Bin, lambdaC, lambdaT, phi, clinicorscreen, caseprop:   REAL;
 BEGIN
  (* Read in the parameter values *)
  p_hi[1,1] := func.arguments[0][0].Value();
  p_hi[1,2] := func.arguments[0][1].Value();
  p_hi[1,3] := func.arguments[0][2].Value();
  p_hi[2,1] := func.arguments[0][3].Value();
  p_hi[2,2] := func.arguments[0][4].Value();
  p_hi[2,3] := func.arguments[0][5].Value();
  p_low[1,1] := func.arguments[0][6].Value();
  p_low[1,2] := func.arguments[0][7].Value();
  p_low[1,3] := func.arguments[0][8].Value();
```

```
p_low[2,1] := func.arguments[0][9].Value();
p_low[2,2] := func.arguments[0][10].Value();
p_low[2,3] := func.arguments[0][11].Value();

Hin := func.arguments[0][12].Value();
Bin := func.arguments[0][13].Value();
lambdaC := func.arguments[0][14].Value();
lambdaT := func.arguments[0][15].Value();
phi := func.arguments[0][16].Value();
clinicorscreen := func.arguments[0][17].Value();
caseprop := func.arguments[0][18].Value();

(* Converts H and B to integer format *)
FOR i:= 1 TO 5000 DO
 IF (Hin = i) THEN;
  H := i;
 END;
END;
FOR i:= 1 TO 5000 DO
 IF (Bin = i) THEN;
  B := i;
 END;
END;

(* Sets B to equal H if follow-up time is shorter than B - Note the program would need
changing if a
   screening study with a follow-up time shorter than B was included *)
IF (H < B) THEN;
 B := H;
END;

(* Calculates the proportion of cases in screening studies infected in the last c = 1 to C days
*)
IF (clinicorscreen = 1) THEN;
 omega_cum[0] := 0;
 FOR b := 1 TO B DO
  omega[b] := phi * (Math.Exp( - lambdaT * (b - 1) / 365) - Math.Exp( - lambdaT * b / 365)) +
              (1 - phi) * (Math.Exp( - lambdaC * (b - 1) / 365) - Math.Exp( - lambdaC * b /
365));                    omega_cum[b] := omega_cum[b-1] + omega[b];
 END;

(* specifies the proportion of women in each state at time zero *)
 pi_hi[0] := caseprop * omega_cum[B];
ELSE
 pi_hi[0] := caseprop;
END;
pi[1,0] := caseprop;
pi[2,0]:= 1 - pi[1,0];
pi[3,0] := 0;
pi_low[0] := pi[1,0] - pi_hi[0];

(* main analysis *)
FOR h := 1 TO B DO;
 pi[1,h] := pi_low[h-1] * p_low[1,1] + pi_hi[h-1] * p_hi[1,1] + pi[2,h-1] * p_hi[2,1];
 pi[2,h] := pi[2,h-1] * p_hi[2,2] + pi_low[h-1] * p_low[1,2] + pi_hi[h-1] * p_hi[1,2];
 pi[3,h]:= 1 - pi[2,h] - pi[1,h];

 IF (clinicorscreen =0) THEN;
  pi_hi[h] := pi[1,h];
  pi_low[h] := 0
```

```
  END;

  IF (clinicorscreen =1) THEN;
   term1[h] := 0;
   FOR i := h TO B DO
    term1[h] := term1[h] + omega[B+1-i] * pi[1,0] * Math.Power(p_hi[1,1],h);
   END;
   term2[1] := 0;
   IF (h >= 2) THEN;
    FOR i := 1 TO h DO
     term2[h] := term2[h-1] + (pi[2,i-1] * p_hi[2,1]) * Math.Power(p_hi[1,1],h-i);
    END;
   END;
   pi_hi[h] := term1[h] + term2[h];
   pi_low[h] := pi[1,h] - pi_hi[h];
  END;
 END;

IF (H > B) THEN;
  FOR h := B+1 TO H DO
   pi[1,h] := pi_low[h-1] * p_low[1,1] + pi_hi[h-1] * p_hi[1,1] + pi[2,h-1] * p_hi[2,1];
   pi[2,h] := pi[2,h-1] * p_hi[2,2] + pi_low[h-1] * p_low[1,2] + pi_hi[h-1] * p_hi[1,2];
   pi[3,h]:= 1 - pi[2,h] - pi[1,h];

   term1[h] := 0;
   FOR i := h+1-B TO h DO
    term2[h] := term2[h-1] + (pi[2,i-1] * p_hi[2,1]) * Math.Power(p_hi[1,1],h-i);
   END;
   pi_hi[h] := term1[h] + term2[h];
   pi_low[h] := pi[1,h] - pi_hi[h];
  END;
 END;
 output := pi[3,H];

 END calculation;



 PROCEDURE (func: Function) Evaluate (OUT value: REAL);
 VAR
  output:                    REAL;

 BEGIN
  calculation(func, output);
  value := output;
 END Evaluate;

 PROCEDURE (f: Factory) New (option: INTEGER): Function;
 VAR
  func: Function;
 BEGIN
  NEW(func); func.Initialize; RETURN func;
 END New;

 PROCEDURE Install*;
 BEGIN
  WBDevScalar.Install(fact);
 END Install;

 PROCEDURE Init;
```

```
VAR
 f: Factory;
BEGIN
 NEW(f); fact := f;
END Init;

BEGIN
 Init;
        END WBDevgeneratep.
```

Appendix 9 Van Valkengoed's method of estimating the probability that a case of *Chlamydia trachomatis* causes pelvic inflammatory disease (see *Chapter 7*)

Van Valkengoed[292] pointed out that the incidence of sequelae of CT, including PID, can provide an upper bound on the progression rate. Subsequently,[39] she proposed a way of using Bayes' theorem to convert information on PID incidence and other items of data into an estimate of the progression probability, i.e. the probability that an episode of CT causes a case of PID. Here we briefly review her approach.

The method begins with the observation that, according to register information in Amsterdam, the incidence of laparoscopically confirmed PID is 34 per 10,000 per year, in women aged 15–40 years. 50% of them would show evidence of previous CT infection, based on serology, but this estimate would provide a highly unreliable estimate of the PEF because of the poor sensitivity of serology tests used at the time and the poor specificity of lifetime exposure to CT as evidence that a PID was caused by CT.

The argument then takes the following form (*Table 47):*

- The incidence of confirmed PID is 34/10,000/year or 0.0034 per year.
- The prevalence of CT is 3%, 0.03.
- The probability of CT in women with PID is 3.84%. This is based on a series of inferences concerning the proportions of women previously exposed to CT, and currently exposed to CT, which is examined below.
- The 3.84% is applied to the 0.0034 incidence of PID, to obtain 0.0001306 in the top left cell in *Table 47*.
- Bayes' theorem is then used:

$$\Pr(PID|CT) = \frac{\Pr(CT|PID).\Pr(PID)}{\Pr(CT)} \tag{83}$$

- In terms of *Table 47*, as a proportion of the CT+ column total 0.03, 0.0001306 is 0.0043 – or 0.43%. This is the final estimate of the probability that a case of incident CT progresses to PID.

TABLE 47 van Valkengoed's method of estimating the proportion of incident CT that progresses to PID

	PID status		
CT status	PID+	PID–	Total
CT+	0.0001306	0.00327	0.0034
CT–	0.029864	0.966136	0.9966
Total	0.03	0.97	1

The 0.43% progression probability is a far lower estimate than had ever been suggested, so low that one might conclude that CT was not a serious public health issue, in the sense that if we multiplied it by our CT incidence estimates from *Chapter 5*, we would predict such a low incidence of PID that screening would be most unlikely to be cost-effective. However, there are a number of problems with the approach. The main difficulty with the argument lies in the derivation of 3.84% as Pr(*CT*|*PID*), which is highly inconsistent with evidence from retrospective studies.[62]

This is based on a further chain of reasoning:

- Serological studies have shown a seroprevalence of 70% in women with current CT infection, and of 26% in currently uninfected women. Assuming perfect sensitivity and specificity, and using serology as a maker of previous or current infection, it was concluded that the proportion of current infection among those who have *ever* been exposed to CT is: $0.70 \times 0.03/(0.70 \times 0.03 + 0.26 \times 0.97) = 0.0769$.
- However, only 50% of PID cases were found to have been ever exposed to CT (as measured by serology), so 0.0769 times $0.5 = 0.0384$, or 3.84% is the proportion of those with clinical PID who are expected to be currently infected with CT.

It is immediately clear that this argument assumes that there is no causal role of CT in PID, and indeed no form of association whatsoever. More precisely, it is assumed that the proportion of current CT among PID cases ever exposed to CT is exactly the same as the proportion of current CT among the total population ever exposed to CT. Estimates of the progression rates from CT to both EP and TFI were also published,[39] based on precisely the same flawed technique.

There are several other aspects of the method that are problematic. The incidence of CT should be used, not its prevalence. Because the duration of CT is in the order of 1 year, however, this does not make much difference to the calculations. The calculations assume that all PID is diagnosed, and, as stated by van Valkengoed,[292] do not account for the poor sensitivity and specificity of CT antibody tests used at the time. Finally, setting aside the difficulties with the reasoning that generated the 3.84 estimate for Pr(*CT*|*PID*), it is important to appreciate that, even if this estimate was credible, the use of Bayes' theorem in this context cannot in and of itself generate a valid estimate of the probability that a CT infection will cause PID. In particular, the 'conditional' symbol 'I ' cannot be interpreted as denoting causation in any way, although if genuinely causal estimates are plugged in, causal inferences from the results would be legitimate. Bayes' theorem as applied to 2×2 tables is simply a method for calculating a row proportion from a column proportion, given knowledge of the marginal totals (see *Table 15*). It implements an equation of the form:

$$\frac{a}{X} = \frac{a}{Y}\frac{Y}{X} \tag{84}$$

Where, for example, a is the top left cell, and X and Y are the proportions of PID and CT, respectively. Nothing in this procedure implies causality or even association. It tells us nothing about the role of CT in the aetiology of PID.

Appendix 10 WinBUGS code for pelvic inflammatory disease incidence synthesis (see *Chapter 7*)

```
model {
# Routine data
for (ag in 1:4) {
 r.routine[ag] ~ dbin(p.routine[ag],N.routine[ag])
 p.routine[ag] <- 1 - exp(-lambda.PID.diag[ag])
 lambda.PID.diag[ag] ~ dexp(0.0001)
 range.temp[ag] ~ dunif(0,range.max[ag])
 range[ag] <- cut(range.temp[ag])
 lambda.PID[ag] <- (lambda.PID.diag[ag] + (range[ag] /
N.routine[ag])) /
                   (1 - psi[1])
# output for Multivariate Normal #prior
 lnlambda.PID[ag] <- log(lambda.PID[ag])
 }
N.routine[1] <- sum(N[16:19])
N.routine[2] <- sum(N[20:24])
N.routine[3] <- sum(N[25:34])
N.routine[4] <- sum(N[35:44])

lambda.PID1624 <- (lambda.PID[1] * N.routine[1] +
                   lambda.PID[2] * N.routine[2]) / sum(N[16:24])
lambda.PID2544 <- (lambda.PID[3] * N.routine[3] +
                   lambda.PID[4] * N.routine[4]) / sum(N[25:44])
lambda.PID1644 <- (lambda.PID[1] * sum(N[16:19]) +
                   lambda.PID[2] * sum(N[20:24]) +
                   lambda.PID[3] * sum(N[25:34]) +
                   lambda.PID[4] * sum(N[35:44])) /
                  sum(N[16:44])
RatioofPIDnums <- (lambda.PID2544 * (N.routine[3] + N.routine[4])) /
                  (lambda.PID1624 * (N.routine[1] + N.routine[2]) +
                   lambda.PID2544 * (N.routine[3] + N.routine[4]))


# Wolner Hanssen
r.wh.undiagpop ~ dbin(psi[1],n.wh.all)
r.wh.asymp ~ dbin(psi[2],n.wh.undiag)
log(lgpsi[1]) <- psi[1]
log(lgpsi[2]) <- psi[2]
logit(lgtpsi[1]) <- psi[1]
logit(lgtpsi[2]) <- psi[2]
psi[1] ~ dbeta(1,1)
psi[2] ~ dbeta(1,1)

# POPI data
r.POPI ~ dbin(p.POPI,n.POPI)
p.POPI <- 1 - exp(-lambda.POPI)
lambda.POPI <- lambda.PID1624 * (1 - (psi[1] * psi[2]))

# Residual Deviance
for (ag in 1:4) {
 dev[ag] <- 2 * (r.routine[ag] * log(r.routine[ag] / (p.routine[ag] *
                 N.routine[ag])) + (N.routine[ag] - r.routine[ag]) *
                 log((N.routine[ag] - r.routine[ag])
                 / (N.routine[ag] - (N.routine[ag] *
p.routine[ag]))))
      }
dev[5] <- 2 * (r.wh.undiagpop * log(r.wh.undiagpop / (psi[1] *
n.wh.all)) +
              (n.wh.all - r.wh.undiagpop) * log((n.wh.all -
r.wh.undiagpop)
               /(n.wh.all - (n.wh.all * psi[1]))))
```

```
dev[6] <- 2 * (r.wh.asymp * log(r.wh.asymp / (psi[2] * n.wh.undiag))
+
               (n.wh.undiag - r.wh.asymp) * log((n.wh.undiag -
r.wh.asymp) /
               (n.wh.undiag - (n.wh.undiag * psi[2]))))
dev[7] <- 2 * (r.POPI * log(r.POPI / (p.POPI * n.POPI)) +
               (n.POPI - r.POPI) * log((n.POPI - r.POPI) /
               (n.POPI - (n.POPI * p.POPI))))
sumdev <- sum(dev[ ])
}

# Data
list(
# Routine data
r.routine = c(8295,13241,18851,11914),
range.max = c(1233,3101,9756,9609),

# Census data - 2002
N=c(NA,NA,NA,NA,NA, NA,NA,NA,NA,NA, NA,NA,NA,NA,NA,
305500,306300,296400,291400,294800,
310100,313900,305600,294700,295000,
304100,317000,329600,349600,370300,
380900,376900,387800,390900,399400,
401200,402600,398700,391900,381900, 370900,356200,349000,343800),

# Wolner-Hansenn
r.wh.undiagpop = 25, n.wh.all = 36
r.wh.asymp = 4, n.wh.undiag = 25

# POPI
r.POPI = 23,n.POPI = 1186
)

# Initial values 1
list(
# population PID incidence
lambda.PID.diag = c(0.01,0.01,0.01,0.01), range.temp = c(1,1,1,1),

# Proportion of PID cases diagnosed
psi = c(0.4,0.5)
)

# Initial values 2
list(
# population PID incidence
lambda.PID.diag = c(0.1,0.1,0.1,0.1), range.temp =
c(1000,1000,1000,1000),

# Proportion of PID cases diagnosed
psi = c(0.1,0.1)
)
```

Appendix 11 WinBUGS code for cumulative pelvic inflammatory disease exposure (see *Chapter 8*)

This Appendix provides the programming code used to estimate the distribution of salpingitis and number of subsequent PID episodes in the population. The appendix is set out in three sections. (A) provides the WBDev functions written in Component Pascal which are called by the WinBUGS code shown in sections B and C. (B) is the WinBUGS code that calculates the numbers of PIDs, CT-related and non-CT-related. (C) is the WinBUGS code that develops the comparisons with the Lund data.

(A). WBDEV code

Note that in appendix A the, philap (the proportion of PID cases that would be diagnosed as salpingitis on laparoscopy) parameter can be removed to generate estimates of the distribution of clinical PID and diagnosed clinical PID. To generate estimates for the distribution of salpingitis it can be included as a multiplier in the exponents for all of the transition probabilities.

```
MODULE WBDevCumulativePIDlap;

 IMPORT
   WBDevVector,
   Math;

 TYPE
   Function = POINTER TO RECORD (WBDevVector.Node) END;
   Factory = POINTER TO RECORD (WBDevVector.Factory) END;

 VAR
   fact-: WBDevVector.Factory;

 PROCEDURE (func: Function) DeclareArgTypes (OUT args: ARRAY OF CHAR);
 BEGIN
   args := "vv";
 END DeclareArgTypes;


 PROCEDURE readindata1 (func: Function; OUT lambda_PID: ARRAY OF REAL);
  VAR
   index,a:                          INTEGER;

 BEGIN
   index := 0;
   FOR a := 1 TO 4 DO
     lambda_PID[a] := func.arguments[0][index].Value();
     INC(index);
   END;
 END readindata1;


PROCEDURE readindata2 (func: Function; OUT psi: REAL);
```

```
BEGIN
 psi := func.arguments[0][4].Value();
END readindata2;


PROCEDURE readindata3 (func: Function; OUT eta1: REAL);
 BEGIN
  eta1 := func.arguments[0][5].Value();
 END readindata3;


PROCEDURE readindata4 (func: Function; OUT EF: ARRAY OF REAL);
 VAR
  index,a:                          INTEGER;

 BEGIN
  index := 6;
  FOR a := 1 TO 4 DO
   EF[a] := func.arguments[0][index].Value();
   INC(index);
  END;
 END readindata4;

PROCEDURE readindata5 (func: Function; OUT philap: REAL);
 BEGIN
  philap := func.arguments[0][10].Value();
 END readindata5;

PROCEDURE readindata6 (func: Function; OUT N: ARRAY OF REAL);
 VAR
  index,a:                          INTEGER;

 BEGIN
  index := 0;
  FOR a := 16 TO 44 DO
   N[a] := func.arguments[1][index].Value();
   INC(index);
  END;
 END readindata6;


PROCEDURE tranratesall (lambda_PID: ARRAY OF REAL; eta1: REAL;
                        OUT lambda_PID2: ARRAY OF ARRAY OF REAL);
 VAR
  a       :                         INTEGER;

 BEGIN
  FOR a := 16 TO 19 DO
   lambda_PID2[1,a] := lambda_PID[1] * 0.85;
   lambda_PID2[2,a] := lambda_PID2[1,a] * eta1;
  END;

  FOR a := 20 TO 24 DO
   lambda_PID2[1,a] := lambda_PID[2] * 0.85;
   lambda_PID2[2,a] := lambda_PID2[1,a] * eta1;
  END;

  FOR a := 25 TO 34 DO
   lambda_PID2[1,a] := lambda_PID[3] * 0.85;
   lambda_PID2[2,a] := lambda_PID2[1,a] * eta1;
```

```
 END;

 FOR a := 35 TO 44 DO
  lambda_PID2[1,a] := lambda_PID[4] * 0.85;
  lambda_PID2[2,a] := lambda_PID2[1,a] * eta1;
 END;
END tranratesall;

PROCEDURE tranratesnct (lambda_PID: ARRAY OF REAL; eta1: REAL; EF: ARRAY OF
REAL;
                        OUT lambda_PID2nct: ARRAY OF ARRAY OF REAL);
 VAR
 a       :                       INTEGER;

 BEGIN
  FOR a := 16 TO 19 DO
   lambda_PID2nct[1,a] := (lambda_PID[1]  - lambda_PID[1] * EF[1]) * 0.85;
   lambda_PID2nct[2,a] := lambda_PID2nct[1,a] * eta1;
  END;

  FOR a := 20 TO 24 DO
   lambda_PID2nct[1,a] := (lambda_PID[2]  - lambda_PID[2] * EF[2]) * 0.85;
   lambda_PID2nct[2,a] := lambda_PID2nct[1,a] * eta1;
  END;

  FOR a := 25 TO 34 DO
   lambda_PID2nct[1,a] := (lambda_PID[3]  - lambda_PID[3] * EF[3]) * 0.85;
   lambda_PID2nct[2,a] := lambda_PID2nct[1,a] * eta1;
  END;

  FOR a := 35 TO 44 DO
   lambda_PID2nct[1,a] := (lambda_PID[4]  - lambda_PID[4] * EF[4]) * 0.85;
   lambda_PID2nct[2,a] := lambda_PID2nct[1,a] * eta1;
  END;
 END tranratesnct;


 PROCEDURE tranproball (lambda_PID2: ARRAY OF ARRAY OF REAL; philap: REAL;
                        OUT p: ARRAY OF ARRAY OF ARRAY OF REAL);
 VAR
 a:                       INTEGER;

 BEGIN
  FOR a := 14 TO 15 DO
   p[1,2,a] := 0;
   p[1,1,a] := 0;

   p[2,2,a] := 0;
   p[2,5,a] := 0;
   p[2,3,a] := 0;

   p[3,3,a] := 0;
   p[3,5,a] := 0;
   p[3,4,a] := 0;

   p[4,5,a] := 0;
   p[4,4,a] := 0;

   p[5,5,a] := 0;
   p[5,8,a] := 0;
```

```
 p[5,6,a] := 0;

 p[6,6,a] := 0;
 p[6,8,a] := 0;
 p[6,7,a] := 0;

 p[7,8,a] := 0;
 p[7,7,a] := 0;

 p[8,8,a] := 0;
END;

FOR a := 16 TO 44 DO
 p[1,2,a] := 1 - Math.Exp(-lambda_PID2[1,a] * philap);
 p[1,1,a] := 1 - p[1,2,a];

 p[2,2,a] := 0;
 p[2,5,a] := 1 - Math.Exp(-lambda_PID2[2,a]);
 p[2,3,a] := 1 - p[2,5,a];

 p[3,3,a] := 0;
 p[3,5,a] := 1 - Math.Exp(-lambda_PID2[2,a]);
 p[3,4,a] := 1 - p[3,5,a];

 p[4,5,a] := 1 - Math.Exp(-lambda_PID2[1,a]);
 p[4,4,a] := 1 - p[4,5,a];

 p[5,5,a] := 0;
 p[5,8,a] := 1 - Math.Exp(-lambda_PID2[2,a]);
 p[5,6,a] := 1 - p[5,8,a];

 p[6,6,a] := 0;
 p[6,8,a] := 1 - Math.Exp(-lambda_PID2[2,a]);
 p[6,7,a] := 1 - p[6,8,a];

 p[7,8,a] := 1 - Math.Exp(-lambda_PID2[1,a]);
 p[7,7,a] := 1 - p[7,8,a];

 p[8,8,a] := 1;
END;
END tranproball;


PROCEDURE tranprobnct (lambda_PID2nct: ARRAY OF ARRAY OF REAL; philap: REAL;
                       OUT pnct: ARRAY OF ARRAY OF ARRAY OF REAL);
VAR
a       :                       INTEGER;

BEGIN
 FOR a := 16 TO 44 DO
  pnct[1,2,a] := 1 - Math.Exp(-lambda_PID2nct[1,a] * philap);
  pnct[1,1,a] := 1 - pnct[1,2,a];

  pnct[2,2,a] := 0;
  pnct[2,5,a] :=  1 - Math.Exp(-lambda_PID2nct[2,a]);
  pnct[2,3,a] :=  1 - pnct[2,5,a];

  pnct[3,3,a] := 0;
  pnct[3,5,a] := 1 - Math.Exp(-lambda_PID2nct[2,a]);
  pnct[3,4,a] := 1 - pnct[3,5,a];
```

```
  pnct[4,5,a] := 1 - Math.Exp(-lambda_PID2nct[1,a]);
  pnct[4,4,a] := 1 - pnct[4,5,a];

  pnct[5,5,a] := 0;
  pnct[5,8,a] := 1 - Math.Exp(-lambda_PID2nct[2,a]);
  pnct[5,6,a] := 1 - pnct[5,8,a];

  pnct[6,6,a] := 0;
  pnct[6,8,a] := 1 - Math.Exp(-lambda_PID2nct[2,a]);
  pnct[6,7,a] := 1 - pnct[6,8,a];

  pnct[7,8,a] := 1 - Math.Exp(-lambda_PID2nct[1,a]);
  pnct[7,7,a] := 1 - pnct[7,8,a];

  pnct[8,8,a] :=1;
 END;
END tranprobnct;


PROCEDURE stateoccpropall (p: ARRAY OF ARRAY OF ARRAY OF REAL;
                           OUT pi: ARRAY OF ARRAY OF REAL);
VAR
a,s:                    INTEGER;

BEGIN
 FOR a := 13 TO 15 DO
  pi[1,a] := 1;
  FOR s := 2 TO 8 DO
   pi[s,a] := 0;
  END;
 END;

 FOR a := 16 TO 44 DO
  pi[1,a] := pi[1,a-1] * p[1,1,a];
  pi[2,a] := pi[1,a-1] * p[1,2,a];
  pi[3,a] := pi[2,a-1] * p[2,3,a];
  pi[4,a] := pi[3,a-1] * p[3,4,a] + pi[4,a-1] * p[4,4,a];
  pi[5,a] := pi[2,a-1] * p[2,5,a] + pi[3,a-1] * p[3,5,a] + pi[4,a-1] * p[4,5,a] + pi[5,a-1] * p[5,5,a];
  pi[6,a] := pi[5,a-1] * p[5,6,a];
  pi[7,a] := pi[6,a-1] * p[6,7,a] + pi[7,a-1] * p[7,7,a];
  pi[8,a] := pi[5,a-1] * p[5,8,a] + pi[6,a-1] * p[6,8,a] + pi[7,a-1] * p[7,8,a] + pi[8,a-1];
 END;

END stateoccpropall;

PROCEDURE stateoccpropnct (pnct: ARRAY OF ARRAY OF ARRAY OF REAL;
                                        OUT pinct: ARRAY OF ARRAY OF REAL);
VAR
a,s:                    INTEGER;

BEGIN
 FOR a := 13 TO 15 DO
  pinct[1,a] := 1;
  FOR s := 2 TO 8 DO
   pinct[s,a] := 0;
  END;
 END;

 FOR a := 16 TO 44 DO
```

```
  pinct[1,a] := pinct[1,a-1] * pnct[1,1,a];
  pinct[2,a] := pinct[1,a-1] * pnct[1,2,a];
  pinct[3,a] := pinct[2,a-1] * pnct[2,3,a];
  pinct[4,a] := pinct[3,a-1] * pnct[3,4,a] + pinct[4,a-1] * pnct[4,4,a];
  pinct[5,a] := pinct[2,a-1] * pnct[2,5,a] + pinct[3,a-1] * pnct[3,5,a] + pinct[4,a-1] * pnct[4,5,a] +
                pinct[5,a-1] * pnct[5,5,a];
  pinct[6,a] := pinct[5,a-1] * pnct[5,6,a];
  pinct[7,a] := pinct[6,a-1] * pnct[6,7,a] + pinct[7,a-1] * pnct[7,7,a];
  pinct[8,a] := pinct[5,a-1] * pnct[5,8,a] + pinct[6,a-1] * pnct[6,8,a] + pinct[7,a-1] * pnct[7,8,a] +
                pinct[8,a-1];
 END;
END stateoccpropnct;


PROCEDURE numberPIDsageall (pi: ARRAY OF ARRAY OF REAL; N: ARRAY OF REAL;
                                OUT PIDnum: ARRAY OF ARRAY OF REAL);
VAR
a:              INTEGER;

BEGIN
 FOR a := 16 TO 44 DO
  PIDnum[1,a] := (pi[2,a] + pi[3,a] + pi[4,a]);
  PIDnum[2,a] := (pi[5,a] + pi[6,a] + pi[7,a]);
  PIDnum[3,a] := pi[8,a];
 END;
END numberPIDsageall;


PROCEDURE numberPIDsagegpall (PIDnum: ARRAY OF ARRAY OF REAL; N: ARRAY OF
REAL;
                                OUT numPIDs: ARRAY OF ARRAY OF REAL);
VAR
n:              INTEGER;

BEGIN
 FOR n := 1 TO 3 DO
  numPIDs[n,1] := (PIDnum[n,16] * N[16] + PIDnum[n,17] * N[17] + PIDnum[n,18] * N[18] +
                   PIDnum[n,19] * N[19]) /
                   (N[16] + N[17] + N[18] + N[19]);
  numPIDs[n,2] := (PIDnum[n,20] * N[20] + PIDnum[n,21] * N[21]  + PIDnum[n,22] * N[22]  +
                   PIDnum[n,23] * N[23] + PIDnum[n,24] * N[24]) /
                   (N[20] + N[21] + N[22] + N[23] + N[24]);
  numPIDs[n,3] := (PIDnum[n,25] * N[25] + PIDnum[n,26] * N[26] + PIDnum[n,27] * N[27]  +
                   PIDnum[n,28] * N[28] + PIDnum[n,29] * N[29] + PIDnum[n,30] * N[30] +
                   PIDnum[n,31] * N[31] + PIDnum[n,32] * N[32] + PIDnum[n,33] * N[33]+
                   PIDnum[n,34] * N[34]) /
                 (N[25] + N[26] + N[27] + N[28] + N[29] + N[30] + N[31] + N[32] + N[33] +
N[34]);
numPIDs[n,4] := (PIDnum[n,35] * N[35] + PIDnum[n,36] * N[36] + PIDnum[n,37] * N[37] +
                 PIDnum[n,38] * N[38] + PIDnum[n,39] * N[39] + PIDnum[n,40] * N[40] +
                 PIDnum[n,41] * N[41] + PIDnum[n,42] * N[42] + PIDnum[n,43] * N[43] +
                 PIDnum[n,44] * N[44]) /
                (N[35] + N[36] + N[37] + N[38] + N[39] + N[40] + N[41] + N[42] + N[43] +
N[44]);
  END;
 END numberPIDsagegpall;


 PROCEDURE numberPIDsobsageall (PIDnum: ARRAY OF ARRAY OF REAL; psi: REAL;
                                OUT PIDnumobs: ARRAY OF ARRAY OF REAL);
```

```
 VAR
 a:                INTEGER;

 BEGIN
  FOR a := 16 TO 44 DO
   PIDnumobs[1,a] :=  PIDnum[1,a] * psi + 2 * PIDnum[2,a] * psi * (1 - psi) +
                      3 * PIDnum[3,a] * psi * (1 - psi) * (1 - psi);
   PIDnumobs[2,a] :=  PIDnum[2,a] * psi * psi + 3 * PIDnum[3,a] * psi * psi * (1 - psi);
   PIDnumobs[3,a] :=  PIDnum[3,a] * psi * psi * psi
  END;
 END numberPIDsobsageall;


PROCEDURE numberPIDsagenct (pinct: ARRAY OF ARRAY OF REAL;  N: ARRAY OF
REAL;
                                OUT PIDnumnct: ARRAY OF ARRAY OF REAL);
 VAR
 a:                INTEGER;

 BEGIN
  FOR a := 16 TO 44 DO
   PIDnumnct[1,a] := (pinct[2,a] + pinct[3,a] + pinct[4,a]);
   PIDnumnct[2,a] := (pinct[5,a] + pinct[6,a] + pinct[7,a]);
   PIDnumnct[3,a] := pinct[8,a];
  END;
 END numberPIDsagenct;


  PROCEDURE numberPIDsagegpnct (PIDnumnct: ARRAY OF ARRAY OF REAL; N:
ARRAY OF
                                      REAL;
                                OUT numPIDsnct: ARRAY OF ARRAY OF REAL);

 VAR
 n:                INTEGER;

 BEGIN
  FOR n := 1 TO 3 DO
   numPIDsnct[n,1] := (PIDnumnct[n,16] * N[16] + PIDnumnct[n,17] * N[17] +
                       PIDnumnct[n,18] * N[18] + PIDnumnct[n,19] * N[19]) /
                      (N[16] + N[17] + N[18] + N[19]);
   numPIDsnct[n,2] := (PIDnumnct[n,20] * N[20] + PIDnumnct[n,21] * N[21] +
                       PIDnumnct[n,22] * N[22] + PIDnumnct[n,23] * N[23] +
                       PIDnumnct[n,24] * N[24]) /
                      (N[20] + N[21] + N[22] + N[23] + N[24]);
   numPIDsnct[n,3] := (PIDnumnct[n,25] * N[25] + PIDnumnct[n,26] * N[26] +
                       PIDnumnct[n,27] * N[27] + PIDnumnct[n,28] * N[28] +
                       PIDnumnct[n,29] * N[29] + PIDnumnct[n,30] * N[30] +
                       PIDnumnct[n,31] * N[31] + PIDnumnct[n,32] * N[32] +
                       PIDnumnct[n,33] * N[33] + PIDnumnct[n,34] * N[34]) /
                     (N[25] + N[26] + N[27] + N[28] + N[29] + N[30] + N[31] + N[32] + N[33] +
N[34]);
   numPIDsnct[n,4] := (PIDnumnct[n,35] * N[35] + PIDnumnct[n,36] * N[36] +
                       PIDnumnct[n,37] * N[37] + PIDnumnct[n,38] * N[38] +
                       PIDnumnct[n,39] * N[39] + PIDnumnct[n,40] * N[40] +
                       PIDnumnct[n,41] * N[41] + PIDnumnct[n,42] * N[32] +
                       PIDnumnct[n,43] * N[43] + PIDnumnct[n,44] * N[44]) /
                     (N[35] + N[36] + N[37] + N[38] + N[39] + N[40] + N[41] + N[42] + N[43] +
N[44]);
   END;
```

```
 END numberPIDsagegpnct;


 PROCEDURE numberPIDsobsageallnct (PIDnumnct: ARRAY OF ARRAY OF REAL; psi:
REAL;
                                   OUT PIDnumobsnct: ARRAY OF ARRAY OF REAL);
 VAR
 a:              INTEGER;

 BEGIN
  FOR a := 16 TO 44 DO
   PIDnumobsnct[1,a] :=  PIDnumnct[1,a] * psi + 2 * PIDnumnct[2,a] * psi * (1 - psi) +
                         3 * PIDnumnct[3,a] * psi * (1 - psi) * (1 - psi);
   PIDnumobsnct[2,a] :=  PIDnumnct[2,a] * psi * psi + 3 * PIDnumnct[3,a] * psi * psi * (1 - psi);
   PIDnumobsnct[3,a] :=  PIDnumnct[3,a] * psi * psi * psi
  END;
 END numberPIDsobsageallnct;


PROCEDURE eventsin8 (pi: ARRAY OF ARRAY OF REAL; p: ARRAY OF ARRAY OF
ARRAY OF
                     REAL; lambda_PID2: ARRAY OF ARRAY OF REAL;  N: ARRAY
OF REAL;
                OUT PIDsin8: ARRAY OF REAL);
 VAR
 a:              INTEGER;
 last2yrs:       ARRAY 45 OF REAL;

 BEGIN
  PIDsin8[14] := 0;
  PIDsin8[15] := 0;
  FOR a := 16 TO 44 DO
   last2yrs[a] := pi[5,a-3] * p[5,8,a-2] + pi[5,a-2] * p[5,8,a-1] +
                  pi[6,a-3] * p[6,8,a-2] + pi[6,a-2] * p[6,8,a-1] +
                  pi[7,a-3] * p[7,8,a-2] + pi[7,a-2] * p[7,8,a-1];

   PIDsin8[a] := last2yrs[a] * ( 1 - Math.Exp(-lambda_PID2[2,a])) +
                 (pi[8,a] - last2yrs[a]) * (1 - Math.Exp(-lambda_PID2[1,a]))
  END;
 END eventsin8;

PROCEDURE eventsin8nct (pinct: ARRAY OF ARRAY OF REAL; pnct: ARRAY OF ARRAY
OF
                        ARRAY OF REAL; lambda_PID2nct: ARRAY OF ARRAY OF
REAL;  N:
                        ARRAY OF REAL;
                   OUT PIDsin8nct: ARRAY OF REAL);
 VAR
 a:              INTEGER;
 last2yrs:       ARRAY 45 OF REAL;

 BEGIN
  PIDsin8nct[14] := 0;
  PIDsin8nct[15] := 0;
  FOR a := 16 TO 44 DO
   last2yrs[a] := pinct[5,a-3] * pnct[5,8,a-2] + pinct[5,a-2] * pnct[5,8,a-1] +
                  pinct[6,a-3] * pnct[6,8,a-2] + pinct[6,a-2] * pnct[6,8,a-1] +
                  pinct[7,a-3] * pnct[7,8,a-2] + pinct[7,a-2] * pnct[7,8,a-1];

   PIDsin8nct[a] := last2yrs[a] * ( 1 - Math.Exp(-lambda_PID2nct[2,a])) +
```

```
                    (pinct[8,a] - last2yrs[a]) * (1 - Math.Exp(-lambda_PID2nct[1,a]));
 END;
END eventsin8nct;


PROCEDURE predincforpop (pi: ARRAY OF ARRAY OF REAL; p: ARRAY OF ARRAY OF
ARRAY
                                OF REAL; PIDsin8: ARRAY OF REAL;
                            OUT PIDinc: ARRAY OF ARRAY OF REAL);
 VAR
  a:              INTEGER;

 BEGIN
  FOR a := 16 TO 44 DO
   PIDinc[1,a] := pi[1,a-1] * p[1,2,a];
   PIDinc[2,a] := pi[2,a-1] * p[2,5,a] + pi[3,a-1] * p[3,5,a] + pi[4,a-1] * p[4,5,a];
   PIDinc[3,a] := pi[5,a-1] * p[5,8,a] + pi[6,a-1] * p[6,8,a] + pi[7,a-1] * p[7,8,a];
   PIDinc[4,a] := PIDsin8[a];
  END;
 END predincforpop;

PROCEDURE predincforpopnct (pinct: ARRAY OF ARRAY OF REAL; pnct: ARRAY OF
ARRAY OF
                                    ARRAY OF REAL; PIDsin8nct: ARRAY OF REAL;
                                OUT PIDincnct: ARRAY OF ARRAY OF REAL);
 VAR
 a:              INTEGER;

 BEGIN
  FOR a := 16 TO 44 DO
   PIDincnct[1,a] := pinct[1,a-1] * pnct[1,2,a];
   PIDincnct[2,a] := pinct[2,a-1] * pnct[2,5,a] + pinct[3,a-1] * pnct[3,5,a] +
                     pinct[4,a-1] * pnct[4,5,a];
   PIDincnct[3,a] := pinct[5,a-1] * pnct[5,8,a] + pinct[6,a-1] * pnct[6,8,a] +
                     pinct[7,a-1] * pnct[7,8,a];
   PIDincnct[4,a] := PIDsin8nct[a];
   END;
 END predincforpopnct;


PROCEDURE predincgpforpop (PIDinc: ARRAY OF ARRAY OF REAL; N: ARRAY OF
REAL;
                                PIDsin8: ARRAY OF REAL;
                            OUT pred_prop_ag: ARRAY OF REAL);
 VAR
 a:              INTEGER;
 pred_pop:       ARRAY 45 OF REAL;
 sumN:           REAL;

BEGIN
  FOR a := 16 TO 44 DO
   pred_pop[a] := (PIDinc[1,a] + PIDinc[2,a] + PIDinc[3,a] + PIDinc[4,a]) *  N[a];
  END;
  pred_prop_ag[1] := (pred_pop[16] + pred_pop[17] + pred_pop[18] + pred_pop[19]) /
                     (N[16] + N[17] + N[18] + N[19]);
  pred_prop_ag[2] := (pred_pop[20] + pred_pop[21] + pred_pop[22] + pred_pop[23] +
pred_pop[24]) /
                     (N[20] + N[21] + N[22] + N[23] + N[24]);
  pred_prop_ag[3] := (pred_pop[25] + pred_pop[26] + pred_pop[27] + pred_pop[28] +
pred_pop[29] +
```

```
                        pred_pop[30] + pred_pop[31] + pred_pop[32] + pred_pop[33] +
pred_pop[34]) /
                        (N[25] + N[26] + N[27] + N[28] + N[29] + N[30] + N[31] + N[32] + N[33]
+ N[34]);
  pred_prop_ag[4] := (pred_pop[35] + pred_pop[36] + pred_pop[37] + pred_pop[38] +
pred_pop[39] +
                        pred_pop[40] + pred_pop[41] + pred_pop[42] + pred_pop[43] +
pred_pop[44]) /
                       (N[35] + N[36] + N[37] + N[38] + N[39] + N[40] + N[41] + N[42] + N[43] +
N[44]);

  pred_prop_ag[0] := 0;
  sumN := 0;
  FOR a := 16 TO 44 DO
   pred_prop_ag[0] := pred_prop_ag[0] + pred_pop[a];
   sumN := sumN + N[a];
  END;
  pred_prop_ag[0] := pred_prop_ag[0] / sumN;
 END predincgpforpop;

PROCEDURE predincgpforpopnct (PIDincnct: ARRAY OF ARRAY OF REAL; N: ARRAY OF
REAL;
                                    PIDsin8nct: ARRAY OF REAL;
                                OUT pred_prop_agnct: ARRAY OF REAL);
 VAR
 a:              INTEGER;
 pred_pop:       ARRAY 45 OF REAL;
 sumN:           REAL;

 BEGIN

  FOR a := 16 TO 44 DO
   pred_pop[a] := (PIDincnct[1,a] + PIDincnct[2,a] + PIDincnct[3,a] + PIDincnct[4,a]) *  N[a];
  END;
  pred_prop_agnct[1] := (pred_pop[16] + pred_pop[17] + pred_pop[18] + pred_pop[19]) /
                        (N[16] + N[17] + N[18] + N[19]);
  pred_prop_agnct[2] := (pred_pop[20] + pred_pop[21] + pred_pop[22] + pred_pop[23] +
                        pred_pop[24])  / (N[20] + N[21] + N[22] + N[23] + N[24]);
  pred_prop_agnct[3] := (pred_pop[25] + pred_pop[26] + pred_pop[27] + pred_pop[28] +
                        pred_pop[29] + pred_pop[30] + pred_pop[31] + pred_pop[32] +
                        pred_pop[33] + pred_pop[34]) /
                       (N[25] + N[26] + N[27] + N[28] + N[29] + N[30] + N[31] + N[32] + N[33]
+ N[34]);
  pred_prop_agnct[4] := (pred_pop[35] + pred_pop[36] + pred_pop[37] + pred_pop[38] +
                        pred_pop[39] + pred_pop[40] + pred_pop[41] + pred_pop[42] +
                        pred_pop[43] + pred_pop[44])   /
                       (N[35] + N[36] + N[37] + N[38] + N[39] + N[40] + N[41] + N[42] + N[43]
+ N[44]);
  pred_prop_agnct[0] := 0;
  sumN := 0;
  FOR a := 16 TO 44 DO
   pred_prop_agnct[0] := pred_prop_agnct[0] + pred_pop[a];
   sumN := sumN + N[a];
  END;
  pred_prop_agnct[0] := pred_prop_agnct[0] / sumN;
 END predincgpforpopnct;


 PROCEDURE (func: Function) Evaluate (OUT values: ARRAY OF REAL);
 VAR
```

```
a, n, i:                                                INTEGER;
eta1, psi, philap:                                      REAL;
lambda_PID, EF, pred_prop_ag, pred_prop_agnct:          ARRAY 5 OF REAL;
PIDsin8,PIDsin8nct,N:                                   ARRAY 45 OF REAL;
lambda_PID2, lambda_PID2nct:                            ARRAY 3,45 OF REAL;
p, pnct:                                                ARRAY 9,9,45 OF REAL;
pi, pinct:                                              ARRAY 9,45 OF REAL;
numPIDs, numPIDsnct:                                    ARRAY 4,5 OF REAL;
PIDinc, PIDincnct:                                      ARRAY 5,45 OF REAL;
PIDnum, PIDnumnct, PIDnumobs,PIDnumobsnct:              ARRAY 4,45 OF REAL;

BEGIN
 readindata1(func, lambda_PID);
 readindata2(func, psi);
 readindata3(func, eta1);
 readindata4(func, EF);
 readindata5(func, philap);
 readindata6(func, N);

 tranratesall(lambda_PID, eta1,   lambda_PID2);
 tranratesnct(lambda_PID, eta1, EF,   lambda_PID2nct);

 tranproball(lambda_PID2, philap,   p);
 tranprobnct(lambda_PID2nct, philap,   pnct);

 stateoccpropall(p,  pi );
 stateoccpropnct(pnct,  pinct );

 numberPIDsageall(pi, N,  PIDnum);
 numberPIDsagenct(pinct, N,   PIDnumnct);

 numberPIDsagegpall(PIDnum, N,  numPIDs);
 numberPIDsagegpnct(PIDnumnct, N,   numPIDsnct);

 numberPIDsobsageall(PIDnum, psi,  PIDnumobs);
 numberPIDsobsageallnct(PIDnumnct, psi,  PIDnumobsnct);

 eventsin8(pi, p, lambda_PID2, N,   PIDsin8);
 predincforpop(pi, p, PIDsin8,  PIDinc);
 predincgpforpop(PIDinc, N, PIDsin8,     pred_prop_ag);

 eventsin8nct(pinct, pnct, lambda_PID2nct, N,   PIDsin8nct);
 predincforpopnct(pinct, pnct, PIDsin8nct,     PIDincnct);
 predincgpforpopnct(PIDincnct, N, PIDsin8nct,       pred_prop_agnct);

 i := 0;
 FOR n := 1TO 3 DO
  FOR a := 1 TO 4 DO
   values[i] := numPIDs[n,a];
   values[i+12] := numPIDsnct[n,a];
   INC(i);
  END;
 END;
 FOR i := 24 TO 28 DO
  values[i] := pred_prop_ag[i-24];
 END;

 i := 29;
 FOR n := 1 TO 3 DO
  FOR a := 16 TO 44 DO
```

```
    values[i] := PIDnum[n,a];
    values[i+87] := PIDnumnct[n,a];
    INC(i);
   END;
  END;
  i := 203;
  FOR n := 1 TO 4 DO
   FOR a := 16 TO 44 DO
    values[i] := PIDinc[n,a];
    values[i+116] := PIDincnct[n,a];
    INC(i);
   END;
  END;
  i := 435;
  FOR n := 1 TO 3 DO
   FOR a := 16 TO 44 DO
    values[i] := PIDnumobs[n,a];
    values[i+87] := PIDnumobsnct[n,a];
    INC(i);
   END;
  END;


 END Evaluate;

 PROCEDURE (f: Factory) New (option: INTEGER): Function;
 VAR
  func: Function;
 BEGIN
  NEW(func); func.Initialize; RETURN func;
 END New;

 PROCEDURE Install*;
 BEGIN
  WBDevVector.Install(fact);
 END Install;

 PROCEDURE Init;
 VAR
  f: Factory;
 BEGIN
  NEW(f); fact := f;
 END Init;

BEGIN
 Init;
END WBDevCumulativePIDlap.
```

(B). WinBUGS code to calculate numbers of PIDs and non-CT related PIDs

```
model {
# PID incidence and ct GP re-infection rate ratio informative priors
Y[1:10] ~ dmnorm(mu[], Omega[ , ])
for (ag in 1:4) {
 log(lambda.PID[ag]) <- Y[ag]
 EF[ag]  <- Y[ag+5] * Y[ag+5]
 }
logit(psi) <- Y[5]
```

```
log(eta1) <- Y[10]

# WBDEV call
for (ag in 1:4) {
 input1[ag] <- lambda.PID[ag]
 input1[ag+6] <- EF[ag]
 }
input1[5] <- psi
input1[6] <- eta1

for (i in 1:29) {
 input2[i] <- N[i+15]
 }

philap ~ dbeta(12,16)
input1[11] <- philap

solution[1:609] <- cumulativePIDlap(input1[1:11],input2[1:29])

for (ag in 1:4) {
 Expect.prop[ag] <- 1 - exp(-lambda.PID[ag])
 }
total.expect.prop <- (Expect.prop[1] * sum(N[16:19]) +
                      Expect.prop[2] * sum(N[20:24]) +
                      Expect.prop[3] * sum(N[25:34]) +
                      Expect.prop[4] * sum(N[35:44])) /
                      sum(N[16:44])
}


# Data
list(
# PID incidence, r-infection raqte, Etological fractions and re-
# infection rate
mu = c(-3.865, -3.595, -3.964, -4.402, -0.5856, 0.7104, 0.4815, 0.3117, 0.3277,
1.919),
Omega = structure(.Data =c(
573.855, -49.264, -10.817, -3.992, 323.371, 97.687, -112.351, -30.198, -5.401, -2.713,
-49.264, 301.419, -14.194, -3.299, 147.591, -78.069, 176.480, -72.284, -11.551, 2.012,
-10.817, -14.194, 105.908, -20.063, 37.684, -13.275, -46.636, 236.305, -118.245,
0.515,
-3.992, -3.299, -20.063, 57.687, 19.172, -2.664, -8.418, -125.677, 138.728, 0.088,
323.371, 147.591, 37.684, 19.172, 351.620, 1.363, -1.443, -1.164, 1.537, 0.030,
97.687, -78.069, -13.275, -2.664, 1.363, 285.706, -324.389, -85.846, -16.864, -7.957,
-112.351, 176.480, -46.636, -8.418, -1.443, -324.389, 734.182, -295.254, -50.046,
8.090,
-30.198, -72.284, 236.305, -125.677, -1.164, -85.846, -295.254, 1512.713, -758.618,
3.088,
-5.401, -11.551, -118.245, 138.728, 1.537, -16.864, -50.046, -758.618, 839.634, 0.705,
-2.713, 2.012, 0.515, 0.088, 0.030, -7.957, 8.090, 3.088, 0.705, 13.711
),
.Dim = c(10,10)),

# Population sizes from census, age =1...44 - 2002
N=c(NA,NA,NA,NA,NA, NA,NA,NA,NA,NA, NA,NA,NA,NA,NA,
305500,306300,296400,291400,294800,
310100,313900,305600,294700,295000,
304100,317000,329600,349600,370300,
380900,376900,387800,390900,399400,
401200,402600,398700,391900,381900, 370900,356200,349000,343800)
)

# Initial values 1
list(
```

```
Y = c(-5,-5,-5,-5,-1,-5,-5,-5,-5,-1)
)

# Initial values 2
list(
Y = c(-1,-1,-1,-1,3,-1,-1,-1,-1,3)
)
```

(C). WinBUGS code comparisons to Lund data

Westrom proportions

```
model {
r1[1:3] ~ dmulti(p1[1:3],N1)
r2[1:3] ~ dmulti(p2[1:3],N2)

p1[1:3] ~ ddirch(a1[1:3])
p2[1:3] ~ ddirch(a2[1:3])
}

# Data
list(r1 = c(771,158,61),r2 = c(220,27,4), N1 = 990, N2 = 251,
     a1 =c(1,1,1), a2 = c(1,1,1))

# Initial Values
list(p1 = c(0.7,0.2,0.1),p2 = c(0.7,0.2,0.1))

# twelve - comparison of cumulative PID to Westrom
model {
 Y[1:10] ~ dmnorm(mu[], Omega[ , ])
 for (ag in 1:4) {
  log(lambda.PID[ag]) <- Y[ag]
  }
 logit(psi) <- Y[5]
 log(eta1) <- Y[10]


# transition rates
 for (a in 16:19) {
  lambda.PID2[1,a] <- lambda.PID[1] * 0.85
  lambda.PID2[2,a] <- lambda.PID2[1,a] * eta1
  }
 for (a in 20:24) {
  lambda.PID2[1,a] <- lambda.PID[2] * 0.85
  lambda.PID2[2,a] <- lambda.PID2[1,a] * eta1
  }
 for (a in 25:34) {
  lambda.PID2[1,a] <- lambda.PID[3] * 0.85
  lambda.PID2[2,a] <- lambda.PID2[1,a] * eta1
 }
 for (a in 35:44) {
  lambda.PID2[1,a] <- lambda.PID[4] * 0.85
  lambda.PID2[2,a] <- lambda.PID2[1,a] * eta1
  }

 # transition probabilities
 # state 1 = 1 PID <1 year
 # state 2 = 1 PID 1-2 years
 # state 3 = 1 PID 2+ years
 # state 4 = 2 PIDs <1 year
```

```
# state 5 = 2 PIDs 1-2 years
# state 6 = 2 PIDs 2+ years
# state 7 = 3 PIDs < 1 year
# state 8 = 3 PIDs 1-2 years
# state 9 = 3 PIDs 2+ years
# state 10 = 4 PIDs < 1 year
# state 11 = 4 PIDs < 1-2 years
# state 12 = 4 PIDs < 2+ years
# state 13 = 5 PIDs assume no more than 5 PIDs

for (a in 16:44) {
 p[1,1,a] <- 0
 p[1,2,a] <- 1 - p[1,4,a]
 p[1,4,a] <- 1 - exp(-lambda.PID2[2,a])

 p[2,2,a] <- 0
 p[2,3,a] <- 1 - p[2,4,a]
 p[2,4,a] <- 1 - exp(-lambda.PID2[2,a])

 p[3,3,a] <- 1 - p[3,4,a]
 p[3,4,a] <- 1 - exp(-lambda.PID2[1,a])

 p[4,4,a] <- 0
 p[4,5,a] <- 1 - p[4,7,a]
 p[4,7,a] <- 1 - exp(-lambda.PID2[2,a])

 p[5,5,a] <- 0
 p[5,6,a] <- 1 - p[5,7,a]
 p[5,7,a] <- 1 - exp(-lambda.PID2[2,a])

 p[6,6,a] <- 1 - p[6,7,a]
 p[6,7,a] <- 1 - exp(-lambda.PID2[1,a])

 p[7,7,a] <- 0
 p[7,8,a] <- 1 - p[7,10,a]
 p[7,10,a] <- 1 - exp(-lambda.PID2[2,a])

 p[8,8,a] <- 0
 p[8,9,a] <- 1 - p[8,10,a]
 p[8,10,a] <- 1 - exp(-lambda.PID2[2,a])

 p[9,9,a] <- 1 - p[9,10,a]
 p[9,10,a] <- 1 - exp(-lambda.PID2[1,a])

 p[10,10,a] <- 0
 p[10,11,a] <- 1 - p[10,13,a]
 p[10,13,a] <- 1 - exp(-lambda.PID2[2,a])

 p[11,11,a] <- 0
 p[11,12,a] <- 1 - p[11,13,a]
 p[11,13,a] <- 1 - exp(-lambda.PID2[2,a])

 p[12,12,a] <- 1 - p[9,10,a]
 p[12,13,a] <- 1 - exp(-lambda.PID2[1,a])

 p[13,13,a] <- 1
}

# State occupancy proportions
for (i in 16:37) {
 pi[i,1,i-1] <- 1
```

```
 pi[i,2,i-1] <- 0
 pi[i,3,i-1] <- 0
 pi[i,4,i-1] <- 0
 pi[i,5,i-1] <- 0
 pi[i,6,i-1] <- 0
 pi[i,7,i-1] <- 0
 pi[i,8,i-1] <- 0
 pi[i,9,i-1] <- 0
 pi[i,10,i-1] <- 0
 pi[i,11,i-1] <- 0
 pi[i,12,i-1] <- 0
 pi[i,13,i-1] <- 0
}

for (i in 16:37) {
 for (a in i:i+7) {
  pi[i,1,a] <- pi[i,1,a-1] * p[1,1,a]
  pi[i,2,a] <- pi[i,1,a-1] * p[1,2,a] + pi[i,2,a-1] * p[2,2,a]
  pi[i,3,a] <- pi[i,2,a-1] * p[2,3,a] + pi[i,3,a-1] * p[3,3,a]
  pi[i,4,a] <- pi[i,1,a-1] * p[1,4,a] + pi[i,2,a-1] * p[2,4,a] +
               pi[i,3,a-1] * p[3,4,a] + pi[i,4,a-1] * p[4,4,a]
  pi[i,5,a] <- pi[i,4,a-1] * p[4,5,a] + pi[i,5,a-1] * p[5,5,a]
  pi[i,6,a] <- pi[i,5,a-1] * p[5,6,a] + pi[i,6,a-1] * p[6,6,a]
  pi[i,7,a] <- pi[i,4,a-1] * p[4,7,a] + pi[i,5,a-1] * p[5,7,a] +
               pi[i,6,a-1] * p[6,7,a] + pi[i,7,a-1] * p[7,7,a]

  pi[i,8,a] <- pi[i,7,a-1] * p[7,8,a] + pi[i,8,a-1] * p[8,8,a]
  pi[i,9,a] <- pi[i,8,a-1] * p[8,9,a] + pi[i,9,a-1] * p[9,9,a]
  pi[i,10,a] <- pi[i,7,a-1] * p[7,10,a] + pi[i,8,a-1] * p[8,10,a] +
                pi[i,9,a-1] * p[9,10,a] + pi[i,10,a-1] * p[10,10,a]
  pi[i,11,a] <- pi[i,10,a-1] * p[10,11,a] + pi[i,11,a-1] *
p[11,11,a]
  pi[i,12,a] <- pi[i,11,a-1] * p[11,12,a] + pi[i,12,a-1] *
p[12,12,a]
  pi[i,13,a] <- pi[i,10,a-1] * p[10,13,a] + pi[i,11,a-1] *
p[11,13,a] +
                pi[i,12,a-1] * p[12,13,a] + pi[i,13,a-1] *
p[13,13,a]
  }
 }

 # Proportion in each state by age
 for (i in 16:37) {
  prop[i,1] <- sum(pi[i,1:3,i+7])
  prop[i,2] <- sum(pi[i,4:6,i+7])
  prop[i,3] <- sum(pi[i,7:9,i+7])
  prop[i,4] <- sum(pi[i,10:12,i+7])
  prop[i,5] <- pi[i,13,i+7]
  propcomp[i,1] <- prop[i,1]
  propcomp[i,2] <- prop[i,2]
  propcomp[i,3] <- sum(prop[i,3:5])
      }
 propunder25[1] <- sum(propcomp[16:24,1]) / 9
 propunder25[2] <- sum(propcomp[16:24,2]) / 9
 propunder25[3] <- sum(propcomp[16:24,3]) / 9
 propover25[1] <- sum(propcomp[25:37,1]) / 13
 propover25[2] <- sum(propcomp[25:37,2]) / 13
 propover25[3] <- sum(propcomp[25:37,3]) / 13

 # Proportion expected to be observed in state by age
 for (i in 16:37) {
```

```
  prop.obs[i,1] <- prop[i,1] +
                   prop[i,2] * (1 - psi) +
                   prop[i,3] * (1 - psi) * (1 - psi) +
                   prop[i,4] * (1 - psi) * (1 - psi) * (1 - psi) +
                  prop[i,5] * (1 - psi) * (1 - psi) * (1 - psi) * (1
- psi)

  prop.obs[i,2] <- prop[i,2] * psi +
                   2 * prop[i,3] * psi * (1 - psi) +
                   3 * prop[i,4] * psi * (1 - psi) * (1 - psi) +
                   4 * prop[i,5] * psi * (1 - psi) * (1 - psi) * (1 -
psi)

  prop.obs[i,3] <- prop[i,3] * psi * psi +
                   3 * prop[i,4] * psi * psi * (1 - psi) +
                   6 * prop[i,5] * psi * psi * (1 - psi) * (1 - psi)

  prop.obs[i,4] <- prop[i,4] * psi * psi * psi  +
                   4 * prop[i,5] * psi * psi * psi * (1 - psi)
  prop.obs[i,5] <- prop[i,5] * psi * psi * psi * psi

  prop.obscomp[i,1] <- prop.obs[i,1]
  prop.obscomp[i,2] <- prop.obs[i,2]
  prop.obscomp[i,3] <- sum(prop.obs[i,3:5])
  }
 prop.obsunder25[1] <- sum(prop.obscomp[16:24,1]) / 9
 prop.obsunder25[2] <- sum(prop.obscomp[16:24,2]) / 9
 prop.obsunder25[3] <- sum(prop.obscomp[16:24,3]) / 9
 prop.obsover25[1] <- sum(prop.obscomp[25:37,1]) / 13
 prop.obsover25[2] <- sum(prop.obscomp[25:37,2]) / 13
 prop.obsover25[3] <- sum(prop.obscomp[25:37,3]) / 13
}


# Data
list(
# PID incidence, r-infection raqte, Etological fractions and re-
# infection rate
mu = c(-3.865, -3.595, -3.964, -4.402, -0.5856, 0.7104, 0.4815, 0.3117, 0.3277,
1.919),
Omega = structure(.Data =c(
573.855, -49.264, -10.817, -3.992, 323.371, 97.687, -112.351, -30.198, -5.401, -2.713,
-49.264, 301.419, -14.194, -3.299, 147.591, -78.069, 176.480, -72.284, -11.551, 2.012,
-10.817, -14.194, 105.908, -20.063, 37.684, -13.275, -46.636, 236.305, -118.245,
0.515,
-3.992, -3.299, -20.063, 57.687, 19.172, -2.664, -8.418, -125.677, 138.728, 0.088,
323.371, 147.591, 37.684, 19.172, 351.620, 1.363, -1.443, -1.164, 1.537, 0.030,
97.687, -78.069, -13.275, -2.664, 1.363, 285.706, -324.389, -85.846, -16.864, -7.957,
-112.351, 176.480, -46.636, -8.418, -1.443, -324.389, 734.182, -295.254, -50.046,
8.090,
-30.198, -72.284, 236.305, -125.677, -1.164, -85.846, -295.254, 1512.713, -758.618,
3.088,
-5.401, -11.551, -118.245, 138.728, 1.537, -16.864, -50.046, -758.618, 839.634, 0.705,
-2.713, 2.012, 0.515, 0.088, 0.030, -7.957, 8.090, 3.088, 0.705, 13.711
),
.Dim = c(10,10))
)


# Initial values 1
list(
Y = c(-5,-5,-5,-5,-1,-5,-5,-5,-5,-1)
)
```

```
# Initial values 2
list(
Y = c(-1,-1,-1,-1,3,-1,-1,-1,-1,3)
)
```

Appendix 12 WinBUGS code for prediction of ectopic pregnancy rates in *Table 35* (see *Chapter 9*)

```
model {
# Population (disecrete time) model - 2 year higher rate
# PID incidence and ct GP re-infection rate ratio informative priors
Y[1:10] ~ dmnorm(mu[], Omega[ , ])
for (ag in 1:4) {
 log(lambda.PID[ag]) <- Y[ag]
 EF[ag] <- Y[ag+5] * Y[ag+5]
 }
logit(psi) <- Y[5]
log(eta1) <- Y[10]

# WBDEV call
for (ag in 1:4) {
 input1[ag] <- lambda.PID[ag]
 input1[ag+6] <- EF[ag]
 }
input1[5] <- psi
input1[6] <- eta1

philap ~ dbeta(12,16)
input1[11] <- philap

for (i in 1:29) {
 input2[i] <- N[i+15]
 }

solution[1:609] <- cumulativePIDlap(input1[1:11],input2[1:29])


# number of previous PIDs [n,a] n = number: 0,1,2,3; a = age-group
for (a in 1:4) {
 numPIDs[1,a] <- 1 - numPIDs[2,a] - numPIDs[3,a] - numPIDs[4,a]
 numPIDs[2,a] <- solution[a]
 numPIDs[3,a] <- solution[a+4]
 numPIDs[4,a] <- solution[a+8]
 numPIDsnct[1,a] <- 1 - numPIDsnct[2,a] - numPIDsnct[3,a] -
numPIDsnct[4,a]
 numPIDsnct[2,a] <- solution[a+12]
 numPIDsnct[3,a] <- solution[a+16]
 numPIDsnct[4,a] <- solution[a+20]
 numPIDsct[1,a] <- numPIDs[1,a] - numPIDsnct[1,a]
 numPIDsct[2,a] <- numPIDs[2,a] - numPIDsnct[2,a]
 numPIDsct[3,a] <- numPIDs[3,a] - numPIDsnct[3,a]
 numPIDsct[4,a] <- numPIDs[4,a] - numPIDsnct[4,a]
 }

# population denominators  - 2002
N1619 <- sum(N[16:19])
N2024 <- sum(N[20:24])
N2534 <- sum(N[25:34])
N3544 <- sum(N[35:44])
N1644 <- sum(N[16:44])
N1617 <- sum(N[16:17])
N1820 <- sum(N[18:20])
N2124 <- sum(N[21:24])
N2529 <- sum(N[25:29])
N3044 <- sum(N[30:44])
N1819 <- sum(N[18:19])
N3034 <- sum(N[30:34])
```

```
# Distribution of PID severity - 2002
# hospital
r.HESPID[1] ~ dbin(HESPID[1],N1619)
r.HESPID[2] ~ dbin(HESPID[2],N2024)
r.HESPID[3] ~ dbin(HESPID[3],N2534)
r.HESPID[4] ~ dbin(HESPID[4],N3544)

# priors
for (a in 1:4) {
 HESPID[a] ~ dbeta(1,1)
 }

# kc60
r.kc602008[1] ~ dbin(kc602008[1],N1619)
r.kc602008[2] ~ dbin(kc602008[2],N2024)
r.kc602008[3] ~ dbin(kc602008[3],N2534)
r.kc602008[4] ~ dbin(kc602008[4],N3544)

# priors
for (a in 1:4) {
 kc602008[a] ~ dbeta(1,1)
 }

# data by age not available until 2008
# 13421 / 12117 is the ratio of totals for 2002 and GUMCAD data for
# 2008
# assumes GUMCAD data representative of all kc60
# assumes age distribution is the same in 2002 as 2008
for (a in 1:4) {
 kc60[a] <- kc602008[a] * 13421 / 12117
 }

# GPRD
r.GPRDPID[1] ~ dbin(GPRDPID[1],N1619)
r.GPRDPID[2] ~ dbin(GPRDPID[2],N2024)
r.GPRDPID[3] ~ dbin(GPRDPID[3],N2534)
r.GPRDPID[4] ~ dbin(GPRDPID[4],N3544)

# priors
for (a in 1:4) {
 GPRDPID[a] ~ dbeta(1,1)
 }

# Distribution
for (a in 1:4) {
 pmin[a] <- kc60[a] + max(HESPID[a],GPRDPID[a])
 pmax[a] <- kc60[a] + HESPID[a] + GPRDPID[a]
 X[a] ~ dunif(pmin[a],pmax[a])
 hospdiag[a] <- psi * HESPID[a] / X[a]
 milddiag[a] <- 1 - hospdiag[a] - undiag[a]
 undiag[a] <- (1 - psi)
 }

# proportion with PIDs by each age, severity, and number
# all PID
# n = 1
for (s in 1:3) {
 for (a in 1:4) {
  PIDcat[1,s,a] <- 1 - (PIDcat[2,s,a] + PIDcat[3,s,a] +
PIDcat[4,s,a])
  }
```

```
 }

# n = 2-4
for (n in 2:4) {
      for (a in 1:4) {
  PIDcat[n,1,a] <- numPIDs[n,a] * undiag[a]
  PIDcat[n,2,a] <- numPIDs[n,a] * milddiag[a]
  PIDcat[n,3,a] <- numPIDs[n,a] * hospdiag[a]
  }
 }

for (n in 2:4) {
 for (s in 1:3) {
  PIDcat1624_2544[n,s,1] <- (PIDcat[n,s,1] * sum(N[16:19]) +
                             PIDcat[n,s,2] * sum(N[20:24]) +
                             PIDcat[n,s,3] * sum(N[25:29])) /
                            sum(N[16:29])
  PIDcat1624_2544[n,s,2] <- (PIDcat[n,s,3] * sum(N[30:34]) +
                             PIDcat[n,s,4] * sum(N[35:44])) /
                            sum(N[30:44])
  PIDcat1624_2544[n,s,3] <- (PIDcat[n,s,1] * sum(N[16:19]) +
                             PIDcat[n,s,2] * sum(N[20:24]) +
                             PIDcat[n,s,3] * sum(N[25:34]) +
                             PIDcat[n,s,4] * sum(N[35:44])) /
                            sum(N[16:44])
  }
 }

for (a in 1:2) {
 for (n in 2:4) {
  sumsevPIDcat1624_2544[n,a] <- sum(PIDcat1624_2544[n, ,a])
  }
 sum2sevPIDcat1624_2544[a] <-sum(sumsevPIDcat1624_2544[2:4,a])
 for (s in 1:3) {
  sumnumPIDcat1624_2544[s,a] <- sum(PIDcat1624_2544[2:4,s,a])
  }
 sum2numPIDcat1624_2544[a] <- sum(sumnumPIDcat1624_2544[ ,a])
 }

# non-CT related PID
# n = 1
for (s in 1:3) {
 for (a in 1:4) {
  PIDcatnct[1,s,a] <- 1 - (PIDcatnct[2,s,a] + PIDcatnct[3,s,a] +
                      PIDcatnct[4,s,a])
  }
 }

# n = 2-4
for (n in 2:4) {
 for (a in 1:4) {
  PIDcatnct[n,1,a] <- numPIDsnct[n,a] * undiag[a]
  PIDcatnct[n,2,a] <- numPIDsnct[n,a] * milddiag[a]
  PIDcatnct[n,3,a] <- numPIDsnct[n,a] * hospdiag[a]
  }
 }

for (n in 2:4) {
 for (s in 1:3) {
  PIDcat1624_2544nct[n,s,1] <- (PIDcatnct[n,s,1] * sum(N[16:19]) +
                                PIDcatnct[n,s,2] * sum(N[20:24]) +
```

```
                                  PIDcatnct[n,s,3] * sum(N[25:29])) /
                                 sum(N[16:29])
   PIDcat1624_2544nct[n,s,2] <- (PIDcatnct[n,s,3] * sum(N[30:34]) +
                                  PIDcatnct[n,s,4] * sum(N[35:44])) /
                                 sum(N[30:44])
   PIDcat1624_2544nct[n,s,3] <- (PIDcatnct[n,s,1] * sum(N[16:19]) +
                                  PIDcatnct[n,s,2] * sum(N[20:24]) +
                                  PIDcatnct[n,s,3] * sum(N[25:34]) +
                                  PIDcatnct[n,s,4] * sum(N[35:44])) /
                                 sum(N[16:44])
   }
  }

for (a in 1:3) {
 for (n in 2:4) {
  sumsevPIDcat1624_2544nct[n,a] <- sum(PIDcat1624_2544nct[n, ,a])
  }
 sum2sevPIDcat1624_2544nct[a] <-sum(sumsevPIDcat1624_2544nct[2:4,a])
 for (s in 1:3) {
  sumnumPIDcat1624_2544nct[s,a] <- sum(PIDcat1624_2544nct[2:4,s,a])
  }
 sum2numPIDcat1624_2544nct[a] <- sum(sumnumPIDcat1624_2544nct[ ,a])
 }

# CT related PID
# n = 1
for (s in 1:3) {
 for (a in 1:4) {
  PIDcatct[1,s,a] <- 1 - (PIDcatct[2,s,a] + PIDcatct[3,s,a] +
                          PIDcatct[4,s,a])
  }
 }

# n = 2-4
for (n in 2:4) {
 for (a in 1:4) {
  PIDcatct[n,1,a] <- numPIDsct[n,a] * undiag[a]
  PIDcatct[n,2,a] <- numPIDsct[n,a] * milddiag[a]
  PIDcatct[n,3,a] <- numPIDsct[n,a] * hospdiag[a]
  }
 }

for (n in 2:4) {
 for (s in 1:3) {
  PIDcat1624_2544ct[n,s,1] <- (PIDcatct[n,s,1] * sum(N[16:19]) +
                               PIDcatct[n,s,2] * sum(N[20:24]) +
                               PIDcatct[n,s,3] * sum(N[25:29])) /
                              sum(N[16:29])
  PIDcat1624_2544ct[n,s,2] <- (PIDcatct[n,s,3] * sum(N[30:34]) +
                               PIDcatct[n,s,4] * sum(N[35:44])) /
                              sum(N[30:44])
  PIDcat1624_2544ct[n,s,3] <- (PIDcatct[n,s,1] * sum(N[16:19]) +
                               PIDcatct[n,s,2] * sum(N[20:24]) +
                               PIDcatct[n,s,3] * sum(N[25:34]) +
                               PIDcatct[n,s,4] * sum(N[35:44])) /
                              sum(N[16:44])

  }
 }

for (a in 1:3) {
```

```
 for (n in 2:4) {
  sumsevPIDcat1624_2544ct[n,a] <- sum(PIDcat1624_2544ct[n, ,a])
  }
 sum2sevPIDcat1624_2544ct[a] <-sum(sumsevPIDcat1624_2544ct[2:4,a])
 for (s in 1:3) {
  sumnumPIDcat1624_2544ct[s,a] <- sum(PIDcat1624_2544ct[2:4,s,a])
  }
 sum2numPIDcat1624_2544ct[a] <- sum(sumnumPIDcat1624_2544ct[ ,a])
 }


# Probability of EP given PID and conception
# Likelihood
# Westrom progression to EP by number of episodes
for (n in 1:4) {
 r.EPnum[n] ~ dbin(EPnum[n],n.EPnum[n])
 }

# Westrom progression to EP by severity in women with 1 PID
for (s in 1:3) {
 r.EPsev[s] ~ dbin(EPsev[s],n.EPsev[s])
 }

# Westrom progression to EP by age in women with 1 PID
for (a in 1:2) {
 r.EPage[a] ~ dbin(EPage[a],n.EPage[a])
 }

# functional retionship between model (see below) parameters and
# likelihood

EPnum[1] <- (control[1] * 713 + control[2] * 199) / 912

for (n in 2:4) {
 EPnum[n] <- (((PIDconctoEP[n,1,1] * n.EPsev[1] +
                PIDconctoEP[n,2,1] * n.EPsev[2] +
                PIDconctoEP[n,3,1] * n.EPsev[3]) /
               sum(n.EPsev[1:3])) * n.EPage[1]
                    +
              ((PIDconctoEP[n,1,2] * n.EPsev[1] +
                PIDconctoEP[n,2,2] * n.EPsev[2] +
                PIDconctoEP[n,3,2] * n.EPsev[3]) /
               sum(n.EPsev[1:3])) * n.EPage[2]) /
             sum(n.EPage[1:2])
 }

for (s in 1:3) {
 EPsev[s] <- (PIDconctoEP[2,s,1] * n.EPage[1] +
              PIDconctoEP[2,s,2] * n.EPage[2]) /
             sum(n.EPage[1:2])
 }

for (a in 1:2) {
 EPage[a] <- (PIDconctoEP[2,1,a] * n.EPsev[1] +
              PIDconctoEP[2,2,a] * n.EPsev[2] +
              PIDconctoEP[2,3,a] * n.EPsev[3]) /
             sum(n.EPsev[1:3])
 }

# probability of EP by age - all PID
```

```
# assume EPs occur 5yrs into Westrom (8.9 year foll) for distribution
# of EP #age
# Progression probabilities by age, severity, and number
# progress[n,s,a] n: number of PIDs 0,1,2,3+, sev mild,mod,sev, a:
# age<=29, #30+
# model
for (s in 1:3) {
 for (a in 1:2) {
  PIDconctoEP[1,s,a] <- EPnum[1] # constant across age!!!
  for (n in 2:4) {
   logit(PIDconctoEP[n,s,a]) <- beta0 + beta1[n] + beta2[s] +
beta3[a]
   }
  }
 }

control[1] <- alphatemp * propEP.1624 / (propEP.1624 + propEP.2544)
control[2] <- alphatemp * propEP.2544 / (propEP.1624 + propEP.2544)

logit(alphatemp) <- alpha

# Priors
# alpha control group thing
alpha ~ dnorm(0,0.0001)

# beta0 = 1 PID, mild, young,
beta0 ~ dnorm(0,0.0001)

# beta1[n] = effect of 0,1,2or3 PIDs beta1[1] = 0, beta1[2] = 0
beta1[1] <- 0
beta1[2] <- 0
for (n in 3:4) {
 beta1[n] ~ dnorm(0,0.0001)
 }

# beta2[s] = effect of severity: beta2[1] = 0
beta2[1] <- 0
for (s in 2:3) {
 beta2[s] ~ dnorm(0,0.0001)
 }

# beta3[a] = effect of age: beta3[1] = 0
beta3[1] <- 0
beta3[2] ~ dnorm(0,0.0001)

# RESIDUAL DEVIANCE
for (m in 1:4) {
 dev1[m] <- 2 * (r.EPnum[m] * log(r.EPnum[m] / (EPnum[m] *
n.EPnum[m])) +
                (n.EPnum[m] - r.EPnum[m]) * log((n.EPnum[m] -
r.EPnum[m]) /
                (n.EPnum[m] - (n.EPnum[m] * EPnum[m]))))
 }

for (s in 1:3) {
 dev2[s] <- 2 * (r.EPsev[s] * log(r.EPsev[s] / (EPsev[s] *
n.EPsev[s])) +
                (n.EPsev[s] - r.EPsev[s]) * log((n.EPsev[s] -
r.EPsev[s]) /
                (n.EPsev[s] - (n.EPsev[s] * EPsev[s]))))
 }
```

```
for (a in 1:2) {
 dev3[a] <- 2 * (r.EPage[a] * log(r.EPage[a] / (EPage[a] *
n.EPage[a])) +
                (n.EPage[a] - r.EPage[a]) * log((n.EPage[a] -
r.EPage[a]) /
                (n.EPage[a] - (n.EPage[a] * EPage[a]))))
 }
sumdev1 <- sum(dev1[])
sumdev2 <- sum(dev2[])
sumdev3 <- sum(dev3[])
sumdev.tot <- sumdev1 + sumdev2 + sumdev3


# Progression probabilities by age, diagnostic status, and number
# progress[n,s,a] n: number of PIDs 0,1,2,3+,
# 1: undiagnosed (mild),
# 2: diagnosed outside of Hospital (mild),
# 3: hospital diagnosed(overall Westrom)
# a: age<=29, 30+

# model 1
#for (n in 2:4) {
# for (a in 1:2) {
#  PIDconctoEP2[n,1,a] <- ((PIDconctoEP[n,1,a] * n.EPsev[1] +
#                           PIDconctoEP[n,2,a] * n.EPsev[2] +
#                           PIDconctoEP[n,3,a] * n.EPsev[3]) /
#                          sum(n.EPsev[1:3])) - control[a]
#  PIDconctoEP2[n,2,a] <- ((PIDconctoEP[n,1,a] * n.EPsev[1] +
#                           PIDconctoEP[n,2,a] * n.EPsev[2] +
#                           PIDconctoEP[n,3,a] * n.EPsev[3]) /
#                          sum(n.EPsev[1:3])) - control[a]
#  PIDconctoEP2[n,3,a] <- ((PIDconctoEP[n,1,a] * n.EPsev[1] +
#                           PIDconctoEP[n,2,a] * n.EPsev[2] +
#                          PIDconctoEP[n,3,a] * n.EPsev[3]) /
#                         sum(n.EPsev[1:3])) - control[a]
#  }
# }

# model 2
#for (n in 2:4) {
# for (a in 1:2) {
#  PIDconctoEP2[n,1,a] <- PIDconctoEP[n,1,a] - control[a]
#  PIDconctoEP2[n,2,a] <- (PIDconctoEP[n,1,a] * n.EPsev[1] +
#                          PIDconctoEP[n,2,a] * n.EPsev[2] +
#                          PIDconctoEP[n,3,a] * n.EPsev[3]) /
#                         sum(n.EPsev[1:3]) - control[a]
#  PIDconctoEP2[n,3,a] <- (PIDconctoEP[n,1,a] * n.EPsev[1] +
#                          PIDconctoEP[n,2,a] * n.EPsev[2] +
#                          PIDconctoEP[n,3,a] * n.EPsev[3]) /
#                         sum(n.EPsev[1:3]) - control[a]
#  }
# }

# model 3
#for (n in 2:4) {
# for (a in 1:2) {
#  PIDconctoEP2[n,1,a] <- PIDconctoEP[n,1,a] - control[a]
#  PIDconctoEP2[n,2,a] <- PIDconctoEP[n,1,a] - control[a]
#  PIDconctoEP2[n,3,a] <- (PIDconctoEP[n,1,a] * n.EPsev[1] +
#                          PIDconctoEP[n,2,a] * n.EPsev[2] +
```

```
#                           PIDconctoEP[n,3,a] * n.EPsev[3]) /
#                          sum(n.EPsev[1:3]) - control[a]
#  }
# }

# model 4
for (n in 2:4) {
 for (a in 1:2) {
  A[n,a] <- PIDconctoEP[n,1,a] - control[a]
  PIDconctoEP2[n,1,a] ~ dunif(0,A[n,a])
  PIDconctoEP2[n,2,a] <- PIDconctoEP[n,1,a] - control[a]
  PIDconctoEP2[n,3,a] <- (PIDconctoEP[n,1,a] * n.EPsev[1] +
                          PIDconctoEP[n,2,a] * n.EPsev[2] +
                          PIDconctoEP[n,3,a] * n.EPsev[3]) /
                         sum(n.EPsev[1:3]) - control[a]
  }
 }


# model 5
#for (n in 2:4) {
# for (a in 1:2) {
#  PIDconctoEP2[n,1,a] <- 0
#  PIDconctoEP2[n,2,a] <- PIDconctoEP[n,1,a] - control[a]
#  PIDconctoEP2[n,3,a] <- (PIDconctoEP[n,1,a] * n.EPsev[1] +
#                          PIDconctoEP[n,2,a] * n.EPsev[2] +
#                          PIDconctoEP[n,3,a] * n.EPsev[3]) /
#                         sum(n.EPsev[1:3]) - control[a]
#  }
# }


# Risk a pregnancy will be ectopic due to PID by age
# all PID

for (n in 2:4) {
 for (s in 1:3) {
  EPs[n,s,1] <- PIDcat[n,s,1] * PIDconctoEP2[n,s,1]
  EPs[n,s,2] <- PIDcat[n,s,2] * PIDconctoEP2[n,s,1]
  EPs[n,s,3] <- PIDcat[n,s,3] * (
               (PIDconctoEP2[n,s,1] * N2529 + PIDconctoEP2[n,s,2] *
N3034)
                / N2534)
  EPs[n,s,4] <- PIDcat[n,s,4] * PIDconctoEP2[n,s,2]
  }
 }
for (a in 1:4) {
 EPbyage[a] <- sum(EPs[2:4 , ,a])
 }
EPbyage.tot <- (EPbyage[1] * sum(N[16:19]) +
                EPbyage[2] * sum(N[20:24]) +
                EPbyage[3] * sum(N[25:34]) +
                EPbyage[4] * sum(N[35:44])) /
               sum(N[16:44])


# non-CT related PID
for (n in 2:4) {
 for (s in 1:3) {
  EPsnct[n,s,1] <- PIDcatnct[n,s,1] * PIDconctoEP2[n,s,1]
  EPsnct[n,s,2] <- PIDcatnct[n,s,2] * PIDconctoEP2[n,s,1]
```

```
  EPsnct[n,s,3] <- PIDcatnct[n,s,3] * (
                (PIDconctoEP2[n,s,1] * N2529 + PIDconctoEP2[n,s,2] *
N3034)
                 / N2534)
  EPsnct[n,s,4] <- PIDcatnct[n,s,4] * PIDconctoEP2[n,s,2]
  }
 }
for (a in 1:4) {
 EPbyagenct[a] <- sum(EPsnct[2:4 , ,a])
 }
EPbyagenct.tot <- (EPbyagenct[1] * sum(N[16:19]) +
                   EPbyagenct[2] * sum(N[20:24]) +
                   EPbyagenct[3] * sum(N[25:34]) +
                   EPbyagenct[4] * sum(N[35:44])) /
                  sum(N[16:44])

for (a in 1:4) {
 EPduetoCTbyage[a] <- EPbyage[a] - EPbyagenct[a]
 }
EPduetoCTbyage.tot <- (EPduetoCTbyage[1] * sum(N[16:19]) +
                       EPduetoCTbyage[2] * sum(N[20:24]) +
                       EPduetoCTbyage[3] * sum(N[25:34]) +
                       EPduetoCTbyage[4] * sum(N[35:44])) /
                      sum(N[16:44])

# proportion of PID related EPs due to CT
for (a in 1:4) {
 propCTofPIDEPs[a] <- 1 - (EPbyagenct[a] / EPbyage[a])
 }
propCTofPIDEPs.tot <- (propCTofPIDEPs[1] * sum(N[16:19]) +
                       propCTofPIDEPs[2] * sum(N[20:24]) +
                       propCTofPIDEPs[3] * sum(N[25:34]) +
                       propCTofPIDEPs[4] * sum(N[35:44])) /
                      sum(N[16:44])

propCTofPIDEPs.tot2 <- (propCTofPIDEPs[1] * r.HESEP[1] +
                        propCTofPIDEPs[2] * r.HESEP[2] +
                        propCTofPIDEPs[3] * r.HESEP[3] +
                        propCTofPIDEPs[4] * r.HESEP[4]) /
                       sum(r.HESEP[1:4])

# Population EP rate by age - 2002
# Conception rate by age - 2002
concdata[1] ~ dbin(conc[1],N1619)
concdata[2] ~ dbin(conc[2],N2024)
concdata[3] ~ dbin(conc[3],N2534)
concdata[4] ~ dbin(conc[4],N3544)

# priors
for (a in 1:4) {
 conc[a] ~ dbeta(1,1)
 }

# EP rate by age - 2002
r.HESEP[1] ~ dbin(HESEP[1],N1619)
r.HESEP[2] ~ dbin(HESEP[2],N2024)
r.HESEP[3] ~ dbin(HESEP[3],N2534)
r.HESEP[4] ~ dbin(HESEP[4],N3544)

# priors
for (a in 1:4) {
```

```
 HESEP[a] ~ dbeta(1,1)
 }

# Analysis of retrospective data
# Odds Ratio
ln.OR.l ~ dnorm(1.22,28.5)
ln.OR.u ~ dnorm(1.69,48.5)
OR.l <- exp(ln.OR.l)
OR.u <- exp(ln.OR.u)


# Population attributable fraction
for (a in 1:4) {
      pi.PID[a] <- 1 - numPIDs[1,a]
      pi.PIDnct[a] <- 1 - numPIDsnct[1,a]
      gamma.l[a] <- (pi.PID[a] * (OR.l - 1)) / (pi.PID[a] * (OR.l -
1) + 1)
      gamma.u[a] <- (pi.PID[a] * (OR.u - 1)) / (pi.PID[a] * (OR.u -
1) + 1)
      gamma.lnct[a] <- (pi.PIDnct[a] * (OR.l - 1)) /
                       (pi.PIDnct[a] * (OR.l - 1) + 1)
      gamma.unct[a] <- (pi.PIDnct[a] * (OR.u - 1)) /
                       (pi.PIDnct[a] * (OR.u - 1) + 1)
      gamma.lct[a] <- gamma.l[a] - gamma.lnct[a]
      gamma.uct[a] <- gamma.u[a] - gamma.unct[a]
      }

pi.PID1624 <- (pi.PID[1] * sum(N[16:19]) + pi.PID[2] * sum(N[20:24]))
/
               sum(N[16:24])
pi.PID2544 <- (pi.PID[3] * sum(N[25:34]) + pi.PID[4] * sum(N[35:44]))
/
               sum(N[25:44])
pi.PID1644 <- (pi.PID[1] * sum(N[16:19]) + pi.PID[2] * sum(N[20:24])
+
               pi.PID[3] * sum(N[25:34]) + pi.PID[4] * sum(N[35:44]))
/
               sum(N[16:44])

pi.PID1624nct <- (pi.PIDnct[1] * sum(N[16:19]) +
                  pi.PIDnct[2] * sum(N[20:24])) / sum(N[16:24])
pi.PID2544nct <- (pi.PIDnct[3] * sum(N[25:34]) +
                  pi.PIDnct[4] * sum(N[35:44])) / sum(N[25:44])
pi.PID1644nct <- (pi.PIDnct[1] * sum(N[16:19]) +
                  pi.PIDnct[2] * sum(N[20:24]) +
                  pi.PIDnct[3] * sum(N[25:34]) +
                  pi.PIDnct[4] * sum(N[35:44])) /
                 sum(N[16:44])

gamma.l.1624 <- (pi.PID1624 * (OR.l - 1)) / (pi.PID1624 * (OR.l - 1)
+ 1)
gamma.u.1624 <- (pi.PID1624 * (OR.u - 1)) / (pi.PID1624 * (OR.u - 1)
+ 1)
gamma.l.2544 <- (pi.PID2544 * (OR.l - 1)) / (pi.PID2544 * (OR.l - 1)
+ 1)
gamma.u.2544 <- (pi.PID2544 * (OR.u - 1)) / (pi.PID2544 * (OR.u - 1)
+ 1)
gamma.l.1644 <- (pi.PID1644 * (OR.l - 1)) / (pi.PID1644 * (OR.l - 1)
+ 1)
gamma.u.1644 <- (pi.PID1644 * (OR.u - 1)) / (pi.PID1644 * (OR.u - 1)
+ 1)
```

```
gamma.l.1624nct <- (pi.PID1624nct * (OR.l - 1)) /
                   (pi.PID1624nct * (OR.l - 1) + 1)
gamma.u.1624nct <- (pi.PID1624nct * (OR.u - 1)) /
                   (pi.PID1624nct * (OR.u - 1) + 1)
gamma.l.2544nct <- (pi.PID2544nct * (OR.l - 1)) /
                   (pi.PID2544nct * (OR.l - 1) + 1)
gamma.u.2544nct <- (pi.PID2544nct * (OR.u - 1)) /
                   (pi.PID2544nct * (OR.u - 1) + 1)
gamma.l.1644nct <- (pi.PID1644nct * (OR.l - 1)) /
                   (pi.PID1644nct * (OR.l - 1) + 1)
gamma.u.1644nct <- (pi.PID1644nct * (OR.u - 1)) /
                   (pi.PID1644nct * (OR.u - 1) + 1)

gamma.l.1624ct <- gamma.l.1624 - gamma.l.1624nct
gamma.u.1624ct <- gamma.u.1624 - gamma.u.1624nct
gamma.l.2544ct <- gamma.l.2544 - gamma.l.2544nct
gamma.u.2544ct <- gamma.u.2544 - gamma.u.2544nct
gamma.l.1644ct <- gamma.l.1644 - gamma.l.1644nct
gamma.u.1644ct <- gamma.u.1644 - gamma.u.1644nct

numEPSduePID.l <- gamma.l.1644 * sum(r.HESEP[ ])
numEPSduePID.u <- gamma.u.1644 * sum(r.HESEP[ ])

numEPSdueCT.l <- gamma.l.1644 * sum(r.HESEP[ ]) * propCTofPIDEPs.tot2
numEPSdueCT.u <- gamma.u.1644 * sum(r.HESEP[ ]) * propCTofPIDEPs.tot2

# Proportion of conceptions that are Ectopic - 2002
for (a in 1:4) {
 propEP[a] <- HESEP[a] / conc[a]
 }
propEP.1624 <- (propEP[1] * sum(N[16:19]) + propEP[2] *
sum(N[20:24])) /
                sum(N[16:24])
propEP.2544 <- (propEP[3] * sum(N[25:34]) + propEP[4] *
sum(N[35:44])) /
                sum(N[25:44])

propEP.tot <- (propEP[1] * sum(N[16:19]) +
               propEP[2] * sum(N[20:24]) +
               propEP[3] * sum(N[25:34]) +
               propEP[4] * sum(N[35:44])) /
              sum(N[16:44])


# Retrospective estimate of prop EPS due to PID - 2002
for (a in 1:4) {
 retroEPfromPID.l[a] <- propEP[a] * gamma.l[a]
 retroEPfromPID.u[a] <- propEP[a] * gamma.u[a]
 }
retroEPfromPID.l.tot <- propEP.tot * gamma.l.1644
retroEPfromPID.u.tot <- propEP.tot * gamma.u.1644

# Proportion due to PID and CT - 2002
for (a in 1:4) {
 propPID[a] <- EPbyage[a] / propEP[a]
 propifnoct[a] <- EPbyagenct[a] / propEP[a]
 propnct[a] <- propPID[a] - propifnoct[a]
 }
propPID.tot <- (propPID[1] * sum(N[16:19]) +
                propPID[2] * sum(N[20:24]) +
```

```
                propPID[3] * sum(N[25:34]) +
                propPID[4] * sum(N[35:44])) /
               sum(N[16:44])
propifnoct.tot <- (propifnoct[1] * sum(N[16:19]) +
                   propifnoct[2] * sum(N[20:24]) +
                   propifnoct[3] * sum(N[25:34]) +
                   propifnoct[4] * sum(N[35:44])) /
                  sum(N[16:44])
propnct.tot <- (propnct[1] * sum(N[16:19]) +
                propnct[2] * sum(N[20:24]) +
                propnct[3] * sum(N[25:34]) +
                propnct[4] * sum(N[35:44])) /
               sum(N[16:44])
}
```

Data

list(

PID incidence, r-infection rate, Etological fractions and re-infection rate

```
mu = c(-3.865, -3.595, -3.964, -4.402, -0.5856, 0.7104, 0.4815, 0.3117, 0.3277,
1.919),
Omega = structure(.Data =c(
573.855, -49.264, -10.817, -3.992, 323.371, 97.687, -112.351, -30.198, -5.401, -2.713,
-49.264, 301.419, -14.194, -3.299, 147.591, -78.069, 176.480, -72.284, -11.551, 2.012,
-10.817, -14.194, 105.908, -20.063, 37.684, -13.275, -46.636, 236.305, -118.245,
0.515,
-3.992, -3.299, -20.063, 57.687, 19.172, -2.664, -8.418, -125.677, 138.728, 0.088,
323.371, 147.591, 37.684, 19.172, 351.620, 1.363, -1.443, -1.164, 1.537, 0.030,
97.687, -78.069, -13.275, -2.664, 1.363, 285.706, -324.389, -85.846, -16.864, -7.957,
-112.351, 176.480, -46.636, -8.418, -1.443, -324.389, 734.182, -295.254, -50.046,
8.090,
-30.198, -72.284, 236.305, -125.677, -1.164, -85.846, -295.254, 1512.713, -758.618,
3.088,
-5.401, -11.551, -118.245, 138.728, 1.537, -16.864, -50.046, -758.618, 839.634, 0.705,
-2.713, 2.012, 0.515, 0.088, 0.030, -7.957, 8.090, 3.088, 0.705, 13.711),
```

.Dim = c(10,10)),

```
r.phi = 26, n.phi = 115,

# Population sizes from census, age =1...44 - 2002
N=c(NA,NA,NA,NA,NA, NA,NA,NA,NA,NA, NA,NA,NA,NA,NA,
305500,306300,296400,291400,294800,
310100,313900,305600,294700,295000,
304100,317000,329600,349600,370300,
380900,376900,387800,390900,399400,
401200,402600,398700,391900,381900, 370900,356200,349000,343800),

# Routine PID data
# HES PID data - 2002
r.HESPID = c(1233,3101,9756,10526), # 16-19, 20-24, 25-34, 35-44

# KC-60 PID data
r.kc602008 = c(2900,3972,3538,1253), #16-19, 20-24, 25-34, 35-44

# predicted GPRD data  - 2002
r.GPRDPID = c(5083,8842,14932,9609), # 16-19, 20-24, 25-34, 35-44

# Prospective PID to EP data (Westrom 560238)
# NB data are for first pregnancy
# results by number of episodes
r.EPnum = c(6,61,24,15),      # 0,1,2,3 diagnosed PIDs
n.EPnum = c(439,912,148,39),  # 0,1,2,3 diagnosed PIDs
```

```
# number by severity, in women with one PID episode only - read from
# graph!
r.EPsev = c(7,19,35),      # mild, moderate, severe
n.EPsev = c(309,420,183),  # mild, moderate, severe

# number by age group, in women with one PID episode only
r.EPage = c(39,22),     # <25, 25-35 at index?
n.EPage = c(713,199),   # <25, 25-35 at index?

# HES EP data - 2002
r.HESEP = c(291,1154,4340,1840), # 16-19, 20-24, 25-34, 35-44


# Conceptions data - 2002
# 13-15, 15-17, 15-19, 20-24, 25-29, 30-34, 35-39, 40+
#conc = c(7900,42000,97100,167800,199400,204300,98900,19600),
# 16-19, 20-24, 25-34, 35-44 - Guessed!
concdata = c(90000,167800,403700,113900)
)


# Initial values 1
list(
Y = c(-5,-5,-5,-5,-1,-5,-5,-5,-5,-1),

HESPID = c(0.01,0.01,0.01,0.01),
kc602008 = c(0.01,0.01,0.01,0.01),
GPRDPID = c(0.01,0.01,0.01,0.01),

alpha = -2,
beta0 = -2,
beta1 = c(NA,NA,0.5,0.5),
beta2 = c(NA,0.05,0.05),
beta3 = c(NA,0.5),

HESEP = c(0.01,0.01,0.01,0.01),
conc = c(0.01,0.01,0.01,0.01),
philap = 0.3
)

# Initial values 2
list(
Y = c(-1,-1,-1,-1,-1,-1,-1,-1,-1,-1),

HESPID = c(0.1,0.1,0.1,0.1),
kc602008 = c(0.1,0.1,0.1,0.1),
GPRDPID = c(0.1,0.1,0.1,0.1),

alpha = -5,
beta0 = -5,
beta1 = c(NA,NA,0.05,0.05),
beta2 = c(NA,0.5,0.5),
beta3 = c(NA,0.05),

HESEP = c(0.1,0.1,0.1,0.1),
conc = c(0.1,0.1,0.1,0.1),
philap = 0.6
)
```

Appendix 13 Overestimation of prevalence of infertility (see *Chapter 10*)

If we took the case of a 34-year-old woman who has been infertile for unknown reasons for 2 years, she would have an approximately 50% chance of conceiving by the age of 44 years.[125] If 50% are sterile, and the remaining 50% could have 0.65 probability of conceiving in any 12-month period, given that they have not yet conceived,[293] then in women who have been infertile for exactly 2 years, one would observe a 56.1% infertility rate, rather than 50%, a 12% multiplicative overestimation.

By the end of 3 years, the overestimation would be down to 4.3%, and by 4 years, only 1.5%. The exact distribution of duration of secondary infertility in surveys of women aged 41–55 years is not known. In estimating the prevalence of secondary infertility in women aged 44 years, we have used a uniform distribution from 0% to 12% to represent this overestimation.

Appendix 14 WinBUGS code and data files for prediction of tubal factor infertility prevalence (see *Chapter 10*)

The following WinBUGS code processes the survey data on infertility (see *Table 37*) and delivers estimates of the prevalence on primary, secondary, and total infertility, and the prevalence of TFI (see *Table 38*). The data input for the Oakley *et al.*[129] study is based on a WinBUGs simulation exercise, shown later in this appendix.

```
Model {
for (i in 1:4) {
 p[i] ~ dbeta(x[1],x[2]) # common beta for pr(1ary infertility)
 }
for (i in 5:7) {
 p[i] ~ dbeta(x[3],x[4]) # common beta for pr(2ndary infertility)
 }
for (i in 8:8) {
 p[i] ~ dbeta(.5,.5)     # Jeffreys priors p[8]
 }
f ~ dunif(0.89,1)        # adjustment for length of follow-up
for (i in 1:4) {
 x[i] ~ dexp(.001)       # priors for beta parameters
 }
for (i in 1:8) {
 r[i] ~ dbin(p[i],n[i])  # likelihood
 rhat[i] <- p[i] * n[i]  # expected value of the numerators
 dev[i] <- 2 * (r[i] * (log(r[i])-log(rhat[i])) + (n[i]-r[i]) *
           (log(n[i]- r[i]) - log(n[i]-rhat[i]))) # Deviance
# contribution
 }
x[5] <- x[1]/sum(x[1:2]) # estimate of pr(1ary infertility)
x[6] <- x[3]/sum(x[3:4]) # estimate of pr(2ndary infertility)
x[7] <- x[6] * f         # adjusted pr(2ndary)
x[8] <- x[5] + x[7]      # total infertility
x[9] <- x[8] * p[8]      #  total TFI
dev[9] <- sum(dev[1:8])  # total residual deviance
dev[10] <- sum(dev[1:7]) # totall res dev for fertility data
}

# Initial Values 1
list(x=c(4,6,4,6,NA,NA,NA,NA),p=c(.4,.4,.4,.4,.4,.4,.4,.4),f=.92))

# Initial values 2
list(x=c(20,20,20,20,NA,NA,NA,NA),p=c(.2,.2,.2,.2,.2,.2,.2,.2),
f=0.96))

# Data
# primary (Bhattacharya, Templeton, Gunnell, Oakley (adjusted))
# secondary ( Bhattacharya, Templeton, Gunnell)
# Proportion of total infertility (including males) due to TFI
# (Maheshwari)
list(r=c(79, 27, 31, 158.3, 5, 17, 41, 442),
n=c(2347, 766, 1609, 6128, 2347, 766, 1609, 1782))
```

Simulation model to adjust the Oakley % primary infertility data for the proportion of women who were involuntarily childless.

```
model {
for (i in 1:2) {
p[i] ~ dbeta(.5,.5)
```

```
r[i] ~ dbin(p[i],n[i]) }
p[3] <-  p[1]/p[2]
}

# Initial values
list(p=c(.5,.5,NA))

# Data
list(r=c(159,2910),n=c(6584,3113))

Results:
```

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
p[1]	0.02422	0.001891	7.665E-6	0.02068	0.02418	0.02806	10001	60000
p[2]	0.9347	0.004425	1.77E-5	0.9257	0.9348	0.943	10001	60000
p[3]	0.02592	0.002027	8.251E-6	0.02211	0.02587	0.03004	10001	60000

The output is used as follows: the probability of primary infertility from the Oakley *et al.*[129] study is taken to be the reported probability divided by the proportion who were not involuntarily childless: (159/6584) / (2910/3113) = 0.02583. The variance of this estimate, taken from the posterior estimate of p[3] is 0.002027^2. The effective denominator is therefore the solution to n in $0.0025834 \times (1 - 0.0025834) / n = 0.002027^2$, which gives $n = 6128$, with numerator 158.3. These figures are used as 'data' in the previous code.

Appendix 15 WinBUGS code for the predictions from prospective models (see *Table 40, Chapter 10*)

```
# TFI analysis
model {
# Population (disecrete time) model - 2 year higher rate
# PID incidence and ct GP re-infection rate ratio informative priors
Y[1:10] ~ dmnorm(mu[], Omega[ , ])
for (ag in 1:4) {
 log(lambda.PID[ag]) <- Y[ag]
 EF[ag] <- Y[ag+5] * Y[ag+5]
 }
logit(psi) <- Y[5]
log(eta1) <- Y[10]

# WBDEV call
for (ag in 1:4) {
 input1[ag] <- lambda.PID[ag]
 input1[ag+6] <- EF[ag]
 }
input1[5] <- psi
input1[6] <- eta1

philap ~ dbeta(12,16)
input1[11] <- philap

for (i in 1:29) {
 input2[i] <- N[i+15]
 }

solution[1:609] <- cumulativePIDlap(input1[1:11],input2[1:29])


# number of previous PIDs [n,a] n = number: 0,1,2,3; a = age-group
for (i in 1:4) {
 numPIDs[2,i] <- solution[i]
 numPIDs[3,i] <- solution[i+4]
 numPIDs[4,i] <- solution[i+8]
 numPIDsnct[2,i] <- solution[i+12]
 numPIDsnct[3,i] <- solution[i+16]
 numPIDsnct[4,i] <- solution[i+20]
 }

# number of previous PIDs in women aged 44, n = number: 0,1,2,3;
num44[2] <- solution[58]
num44[3] <- solution[87]
num44[4] <- solution[116]

num44nct[2] <- solution[145]
num44nct[3] <- solution[174]
num44nct[4] <- solution[203]



# population denominators  - 2002
N1619 <- sum(N[16:19])
N2024 <- sum(N[20:24])
N2534 <- sum(N[25:34])
N3544 <- sum(N[35:44])
N1644 <- sum(N[16:44])
N1617 <- sum(N[16:17])
N1820 <- sum(N[18:20])
N2124 <- sum(N[21:24])
N2529 <- sum(N[25:29])
```

```
N3044 <- sum(N[30:44])
N1819 <- sum(N[18:19])
N3034 <- sum(N[30:34])


# Distribution of PID severity - 2002
# hospital
r.HESPID[1] ~ dbin(HESPID[1],N1619)
r.HESPID[2] ~ dbin(HESPID[2],N2024)
r.HESPID[3] ~ dbin(HESPID[3],N2534)
r.HESPID[4] ~ dbin(HESPID[4],N3544)

# priors
for (a in 1:4) {
 HESPID[a] ~ dbeta(1,1)
 }

# kc60
r.kc602008[1] ~ dbin(kc602008[1],N1619)
r.kc602008[2] ~ dbin(kc602008[2],N2024)
r.kc602008[3] ~ dbin(kc602008[3],N2534)
r.kc602008[4] ~ dbin(kc602008[4],N3544)

# priors
for (a in 1:4) {
 kc602008[a] ~ dbeta(1,1)
 }

# data by age not available until 2008
# 13421 / 12117 is the ratio of totals for 2002 and GUMCAD data for
# 2008
# assumes GUMCAD data representative of all kc60
# assumes age distribution is the same in 2002 as 2008
for (a in 1:4) {
 kc60[a] <- kc602008[a] * 13421 / 12117
 }

# GPRD
r.GPRDPID[1] ~ dbin(GPRDPID[1],N1619)
r.GPRDPID[2] ~ dbin(GPRDPID[2],N2024)
r.GPRDPID[3] ~ dbin(GPRDPID[3],N2534)
r.GPRDPID[4] ~ dbin(GPRDPID[4],N3544)

# priors
for (a in 1:4) {
 GPRDPID[a] ~ dbeta(1,1)
 }

# Distribution
for (a in 1:4) {
pmin[a] <- kc60[a] + max(HESPID[a],GPRDPID[a])
pmax[a] <- kc60[a] + HESPID[a] + GPRDPID[a]
X[a] ~ dunif(pmin[a],pmax[a])
      hospdiag[a] <- psi * HESPID[a] / X[a]
      milddiag[a] <- 1 - hospdiag[a] - undiag[a]
      undiag[a] <- (1 - psi)
      }


# proportion with PIDs by each age, severity, and number
# all PID
```

```
# n = 1
for (s in 1:3) {
 for (a in 1:4) {
  PIDcat[1,s,a] <- 1 - (PIDcat[2,s,a] + PIDcat[3,s,a] +
PIDcat[4,s,a])
  }
 PIDcat44[1,s] <- 1 - (PIDcat44[2,s] + PIDcat44[3,s] + PIDcat44[4,s])

 }

# n = 2-4
for (n in 2:4) {
 for (a in 1:4) {
  PIDcat[n,1,a] <- numPIDs[n,a] * undiag[a]
  PIDcat[n,2,a] <- numPIDs[n,a] * milddiag[a]
  PIDcat[n,3,a] <- numPIDs[n,a] * hospdiag[a]
  }
 PIDcat44[n,1] <- num44[n] * undiag[4]
 PIDcat44[n,2] <- num44[n] * milddiag[4]
 PIDcat44[n,3] <- num44[n] * hospdiag[4]
 }


# non-CT related PID
# n = 1
for (s in 1:3) {
 for (a in 1:4) {
  PIDcatnct[1,s,a] <- 1 - (PIDcatnct[2,s,a] + PIDcatnct[3,s,a] +
                           PIDcatnct[4,s,a])
  }
 PIDcat44nct[1,s] <- 1 - (PIDcat44nct[2,s] + PIDcat44nct[3,s] +
                          PIDcat44nct[4,s])
 }

# n = 2-4
for (n in 2:4) {
 for (a in 1:4) {
  PIDcatnct[n,1,a] <- numPIDsnct[n,a] * undiag[a]
  PIDcatnct[n,2,a] <- numPIDsnct[n,a] * milddiag[a]
  PIDcatnct[n,3,a] <- numPIDsnct[n,a] * hospdiag[a]
  }
 PIDcat44nct[n,1] <- num44nct[n] * undiag[4]
 PIDcat44nct[n,2] <- num44nct[n] * milddiag[4]
 PIDcat44nct[n,3] <- num44nct[n] * hospdiag[4]
 }

# Probability of TFI given PID by age, severity, number of PIDs-
# Westrom
# Likelihood
for (a in 1:2) {
 for (s in 1:3) {
  onePIDsev[a,s] <- p[a,s,1]
  r.onePIDsev[a,s] ~ dbin(onePIDsev[a,s],n.onePIDsev[a,s])
  }
 twoPID[a] <- (p[a,1,2] * n.onePIDsev[a,1] +
               p[a,2,2] * n.onePIDsev[a,2] +
               p[a,3,2] * n.onePIDsev[a,3]) / sum(n.onePIDsev[a, ])
 r.twoPID[a] ~ dbin(twoPID[a],n.twoPID[a])
 threePID[a] <- (p[a,1,3] * n.onePIDsev[a,1] +
                 p[a,2,3] * n.onePIDsev[a,2] +
                 p[a,3,3] * n.onePIDsev[a,3]) / sum(n.onePIDsev[a, ])
```

```
 r.threePID[a] ~ dbin(threePID[a],n.threePID[a])
 }

# model
for (a in 1:2) {
 for (s in 1:3) {
  for (n in 1:3) {
   logit(p[a,s,n]) <- beta0 + beta1[a] + beta2[s] + beta3[n]
   PIDtoTFI[n+1,a,s] <- p[a,s,n]
   }
  }
 }

for (a in 1:2) {
 onePID[a] <- (onePIDsev[a,1] * n.onePIDsev[a,1] +
               onePIDsev[a,2] * n.onePIDsev[a,2] +
               onePIDsev[a,3] * n.onePIDsev[a,3]) /
              sum(n.onePIDsev[a, ])
 }

# priors
beta0 ~ dnorm(0,0.0001)

beta1[1] <- 0
beta1[2] ~ dnorm(0,0.0001)

beta2[1] <- 0
beta2[2] ~ dnorm(0,0.0001)
beta2[3] ~ dnorm(0,0.0001)

beta3[1] <- 0
beta3[2] ~ dnorm(0,0.0001)
beta3[3] ~ dnorm(0,0.0001)

# RESIDUAL DEVIANCE
for (a in 1:2) {
 for (s in 1:3) {
  dev1[a,s] <- 2 * (r.onePIDsev[a,s] * log(r.onePIDsev[a,s] /
                   (onePIDsev[a,s] * n.onePIDsev[a,s])) +
                   (n.onePIDsev[a,s] - r.onePIDsev[a,s]) *
                   log((n.onePIDsev[a,s] - r.onePIDsev[a,s]) /
                   (n.onePIDsev[a,s] - (n.onePIDsev[a,s] *
                    onePIDsev[a,s]))))
  }

 dev2[a] <- 2 * (r.twoPID[a] * log(r.twoPID[a] /
                (twoPID[a] * n.twoPID[a])) +
                (n.twoPID[a] - r.twoPID[a]) *
                 log((n.twoPID[a] - r.twoPID[a]) /
                     (n.twoPID[a] - (n.twoPID[a] * twoPID[a]))))

 dev3[a] <- 2 * (r.threePID[a] * log(r.threePID[a] /
                (threePID[a] * n.threePID[a])) +
                (n.threePID[a] - r.threePID[a]) *
                 log((n.threePID[a] - r.threePID[a]) /
                     (n.threePID[a] - (n.threePID[a] *
threePID[a]))))
 }
sumdev1 <- sum(dev1[ , ])
sumdev2 <- sum(dev2[])
sumdev3 <- sum(dev3[])
```

```
sumdev.tot <- sumdev1 + sumdev2 + sumdev3


# Progression probabilities by age, diagnostic status, and number
# progress[n,s,a] n: number of PIDs 0,1,2,3+,
# 1: undiagnosed (mild),
# 2: diagnosed outside of Hospital (mild),
# 3: hospital diagnosed(overall Westrom)
# a: age<=29, 30+

# model 1
for (n in 1:3) {
 for (a in 1:2) {
      PIDtoTFI2[n,1,a] <- (PIDtoTFI[2,a,1] * n.onePIDsev[a,1] +
                           PIDtoTFI[2,a,2] * n.onePIDsev[a,2] +
                           PIDtoTFI[2,a,3] * n.onePIDsev[a,3]) /
                          sum(n.onePIDsev[a, ])
      PIDtoTFI2[n,2,a] <- (PIDtoTFI[2,a,1] * n.onePIDsev[a,1] +
                           PIDtoTFI[2,a,2] * n.onePIDsev[a,2] +
                           PIDtoTFI[2,a,3] * n.onePIDsev[a,3]) /
                          sum(n.onePIDsev[a, ])
      PIDtoTFI2[n,3,a] <- (PIDtoTFI[2,a,1] * n.onePIDsev[a,1] +
                           PIDtoTFI[2,a,2] * n.onePIDsev[a,2] +
                           PIDtoTFI[2,a,3] * n.onePIDsev[a,3]) /
                          sum(n.onePIDsev[a, ])
           }
      }

# model 2
#for (n in 1:3) {
# for (a in 1:2) {
#  PIDtoTFI2[n,1,a] <- PIDtoTFI[2,a,1]
#  PIDtoTFI2[n,2,a] <- (PIDtoTFI[2,a,1] * n.onePIDsev[a,1] +
#                       PIDtoTFI[2,a,2] * n.onePIDsev[a,2] +
#                       PIDtoTFI[2,a,3] * n.onePIDsev[a,3]) /
#                      sum(n.onePIDsev[a, ])
#  PIDtoTFI2[n,3,a] <- (PIDtoTFI[2,a,1] * n.onePIDsev[a,1] +
#                       PIDtoTFI[2,a,2] * n.onePIDsev[a,2] +
#                       PIDtoTFI[2,a,3] * n.onePIDsev[a,3]) /
#                      sum(n.onePIDsev[a, ])
#  }
# }


# model 3
#for (n in 1:3) {
# for (a in 1:2) {
#  PIDtoTFI2[n,1,a] <- PIDtoTFI[2,a,1]
#  PIDtoTFI2[n,2,a] <- PIDtoTFI[2,a,1]
#  PIDtoTFI2[n,3,a] <- (PIDtoTFI[2,a,1] * n.onePIDsev[a,1] +
#                       PIDtoTFI[2,a,2] * n.onePIDsev[a,2] +
#                       PIDtoTFI[2,a,3] * n.onePIDsev[a,3]) /
#                      sum(n.onePIDsev[a, ])
#  }
# }


# model 4
#for (n in 1:3) {
# for (a in 1:2) {
#  A[n,a] ~ dunif(0,PIDtoTFI[2,a,1])
```

```
#  PIDtoTFI2[n,1,a] <- A[n,a]
#  PIDtoTFI2[n,2,a] <- PIDtoTFI[2,a,1]
#  PIDtoTFI2[n,3,a] <- (PIDtoTFI[2,a,1] * n.onePIDsev[a,1] +
#                       PIDtoTFI[2,a,2] * n.onePIDsev[a,2] +
#                       PIDtoTFI[2,a,3] * n.onePIDsev[a,3]) /
#                      sum(n.onePIDsev[a, ])
#  }
# }


# model 5
#for (n in 1:3) {
# for (a in 1:2) {
#  PIDtoTFI2[n,1,a] <- 0
#  PIDtoTFI2[n,2,a] <- PIDtoTFI[2,a,1]
#  PIDtoTFI2[n,3,a] <- (PIDtoTFI[2,a,1] * n.onePIDsev[a,1] +
#                       PIDtoTFI[2,a,2] * n.onePIDsev[a,2] +
#                       PIDtoTFI[2,a,3] * n.onePIDsev[a,3]) /
#                      sum(n.onePIDsev[a, ])
#  }
# }

# TFIs by age
# all PID
for (n in 2:4) {
 for (s in 1:3) {
  TFIs[n,s,1] <- PIDcat[n,s,1] * PIDtoTFI2[n-1,s,1]
  TFIs[n,s,2] <- PIDcat[n,s,2] * PIDtoTFI2[n-1,s,1]
  TFIs[n,s,3] <- PIDcat[n,s,3] *
                (PIDtoTFI2[n-1,s,1] * N2529 + PIDtoTFI2[n-1,s,2] *
N3034) /
                 N2534
  TFIs[n,s,4] <- PIDcat[n,s,4] * PIDtoTFI2[n-1,s,2]
  TFI44[n,s] <- PIDcat44[n,s] * PIDtoTFI2[n-1,s,2]
  }
 }
for (a in 1:4) {
 TFIbyage[a] <- sum(TFIs[2:4 , ,a])
 }
TFI44tot <- sum(TFI44[2:4 , ])
TFIbyage.tot <- (TFIbyage[1] * sum(N[16:19]) +
                 TFIbyage[2] * sum(N[20:24]) +
                 TFIbyage[3] * sum(N[25:34]) +
                 TFIbyage[4] * sum(N[35:44])) /
                sum(N[16:44])

# non-CT related PID
for (n in 2:4) {
 for (s in 1:3) {
  TFIsnct[n,s,1] <- PIDcatnct[n,s,1] * PIDtoTFI2[n-1,s,1]
  TFIsnct[n,s,2] <- PIDcatnct[n,s,2] * PIDtoTFI2[n-1,s,1]
  TFIsnct[n,s,3] <- PIDcatnct[n,s,3] *
                   (PIDtoTFI2[n-1,s,1] * N2529 + PIDtoTFI2[n-1,s,2] *
                    N3034) / N2534
  TFIsnct[n,s,4] <- PIDcatnct[n,s,4] * PIDtoTFI2[n-1,s,2]
  TFI44nct[n,s] <- PIDcat44nct[n,s] * PIDtoTFI2[n-1,s,2]
  }
 }

for (a in 1:4) {
 TFIbyagenct[a] <- sum(TFIsnct[2:4 , ,a])
```

```
 }
TFI44totnct <- sum(TFI44nct[2:4 , ])
TFIbyagenct.tot <- (TFIbyagenct[1] * sum(N[16:19]) +
                    TFIbyagenct[2] * sum(N[20:24]) +
                    TFIbyagenct[3] * sum(N[25:34]) +
                    TFIbyagenct[4] * sum(N[35:44])) /
                   sum(N[16:44])

for (a in 1:4) {
 TFIduetoCTbyage[a] <- TFIbyage[a] - TFIbyagenct[a]
 }
TFIduetoCTbyage.tot <- (TFIduetoCTbyage[1] * sum(N[16:19]) +
                        TFIduetoCTbyage[2] * sum(N[20:24]) +
                        TFIduetoCTbyage[3] * sum(N[25:34]) +
                        TFIduetoCTbyage[4] * sum(N[35:44])) /
                       sum(N[16:44])

# proportion of PID related TFIs due to CT
for (a in 1:4) {
 propCTofPIDTFIs[a] <- 1 - (TFIbyagenct[a] / TFIbyage[a])
 }
propCTofPIDTFIs44 <- 1 - (TFI44totnct / TFI44tot)
propCTofPIDTFIs.tot <- (propCTofPIDTFIs[1] * sum(N[16:19]) +
                        propCTofPIDTFIs[2] * sum(N[20:24]) +
                        propCTofPIDTFIs[3] * sum(N[25:34]) +
                        propCTofPIDTFIs[4] * sum(N[35:44])) /
                       sum(N[16:44])

}


# Data
list(
# PID incidence, r-infection rate, Etological fractions and re-
infection #rate
mu = c(-3.865, -3.595, -3.964, -4.402, -0.5856, 0.7104, 0.4815, 0.3117, 0.3277,
1.919),
Omega = structure(.Data =c(
573.855, -49.264, -10.817, -3.992, 323.371, 97.687, -112.351, -30.198, -5.401, -2.713,
-49.264, 301.419, -14.194, -3.299, 147.591, -78.069, 176.480, -72.284, -11.551, 2.012,
-10.817, -14.194, 105.908, -20.063, 37.684, -13.275, -46.636, 236.305, -118.245,
0.515,
-3.992, -3.299, -20.063, 57.687, 19.172, -2.664, -8.418, -125.677, 138.728, 0.088,
323.371, 147.591, 37.684, 19.172, 351.620, 1.363, -1.443, -1.164, 1.537, 0.030,
97.687, -78.069, -13.275, -2.664, 1.363, 285.706, -324.389, -85.846, -16.864, -7.957,
-112.351, 176.480, -46.636, -8.418, -1.443, -324.389, 734.182, -295.254, -50.046,
8.090,
-30.198, -72.284, 236.305, -125.677, -1.164, -85.846, -295.254, 1512.713, -758.618,
3.088,
-5.401, -11.551, -118.245, 138.728, 1.537, -16.864, -50.046, -758.618, 839.634, 0.705,
-2.713, 2.012, 0.515, 0.088, 0.030, -7.957, 8.090, 3.088, 0.705, 13.711),
.Dim = c(10,10)),


# Population sizes from census, age =1...44 - 2002
N=c(NA,NA,NA,NA,NA, NA,NA,NA,NA,NA, NA,NA,NA,NA,NA,
305500,306300,296400,291400,294800,
310100,313900,305600,294700,295000,
304100,317000,329600,349600,370300,
380900,376900,387800,390900,399400,
401200,402600,398700,391900,381900, 370900,356200,349000,343800),

# Routine PID data
# HES PID data - 2002
```

```
r.HESPID = c(1233,3101,9756,10526), # 16-19, 20-24, 25-34, 35-44

# KC-60 PID data
r.kc602008 = c(2900,3972,3538,1253), #16-19, 20-24, 25-34, 35-44

# predicted GPRD data  - 2002
r.GPRDPID = c(5083,8842,14932,9609), # 16-19, 20-24, 25-34, 35-44

# Westrom Progression data (A cohort study of 1,844)-assumes a=1 goes
# to 29
r.onePIDsev = structure(.Data =c(
2,23,34,
0,5,15
),
.Dim = c(2,3)),

n.onePIDsev = structure(.Data =c(
241,361,169,
71,89,60
),
.Dim = c(2,3)),

r.twoPID = c(29,7), n.twoPID = c(158,27),
r.threePID = c(23,3), n.threePID = c(61,4),
# Note: 0 TFIs in 451 control women so assume zero.
)

# Initial values 1
list(
Y = c(-5,-5,-5,-5,-1,-5,-5,-5,-5,-1),

HESPID = c(0.01,0.01,0.01,0.01),
kc602008 = c(0.01,0.01,0.01,0.01),
GPRDPID = c(0.01,0.01,0.01,0.01),

philap = 0.3,

beta0 = -2,
beta1 = c(NA,0.5),
beta2 = c(NA,0.05,0.05),
beta3 = c(NA,0.5,0.5)
)
A = structure(.Data =c(
0.01,0.01,
0.01,0.01,
0.01,0.01
),
.Dim = c(3,2)),
)

# Initial values 2
list(
Y = c(-1,-1,-1,-1,-1,-1,-1,-1,-1,-1),

HESPID = c(0.1,0.1,0.1,0.1),
kc602008 = c(0.1,0.1,0.1,0.1),
GPRDPID = c(0.1,0.1,0.1,0.1),

philap = 0.6,

beta0 = -5,
```

```
beta1 = c(NA,0.05),
beta2 = c(NA,0.5,0.5),
beta3 = c(NA,0.05,0.05),
)

A = structure(.Data =c(
0.1,0.1,
0.1,0.1,
0.1,0.1
),
.Dim = c(3,2)),
)
```

Appendix 16 Sensitivity, specificity and resolution of assays (see *Chapter 11*)

Analyses of the Wills *et al.*,[269] Morre *et al.*[265] and Narvanen *et al.*[268] data on test performance, according to our model for Test Resolution, are set out in *Table 48*.

Wills *et al.*[269] examined four assays in 164 female NAAT-positive patients attending STD clinics who had been infected for at least 1 month, and in 747 children aged 2–13 years. Twenty-five of the 747 paediatric samples were positive on MIF, and seven of these were positive on Western blot and were excluded. In our analysis, we assign the remaining 18 MIF-positive samples a uniform prior probability of between zero and one as being true positive. The estimates of test resolution varied from 3.0 to 4.9 (mean = 4.1, SD = 0.83).

Morre *et al.*[265] describe a 'discrepancy analysis' of MIF and three peptide-based assays in 149 STI clinic attenders identified through screening. Women were counted as 'true positive' if they were positive on two peptide assays, or MIF on two occasions. 'Grey' samples were included in our analysis and assigned a uniform prior probability of being true positive between zero and one. Values of R range from 4.7 to 6.4 (mean = 5.9, SD = 0.79).

Narvanen *et al.*[268] tested 82 culture-positive infected women and 152 paediatric control subjects with a CT enzyme immunoassay (CT-EIA). Resolution was 6.29.

TABLE 48 Estimates of sensitivity, false-positive rate and resolution R_T from serological studies

Study	Assay	TP	FN	TN(a)	FP(a)	TN(b)	FP(b)	Grey	Se_T% (95% CI)	Fp_T% (95% CI)	R_T
Morre[265]	CT-EIA	61	11	72	1			4	82.5 (72.7 to 90.7)	1.4 (0 –5)	6.4 (4.4 to 9.4)
	SeroCT	61	11	69	1			7	81.2 (70.6 to 90.2)	1.4 (0 to 5)	6.3 (4.3 to 9.3)
	CT pELISA	50	20	72	2			5	69.3 (57.9 to 79.6)	2.6 (0.3 to 7)	4.7 (3.3 to 6.6)
	MIF	57	15	74	1			2	78.1 (67.9 to 86.9)	1.4 (0 to 4.9)	6.1 (4.2 to 9.2)
Wills[269]	Pgp3	121	43	29	711	17	705		73.8 (66.8 to 80.2)	3.1 (1.9 to 4.6)	4.5 (3.9 to 5.1
	Ani labs	98	66	8	587	6	574		59.8 (52.1 to 67.3)	1.2 (0.6 to 2.0)	4.9 (4.2 to 5.6)
	SeroCT	91	73	22	718	20	702		55.5 (47.8 to 63.0)	2.9 (2.0 to 3.9)	3.8 (3.3 to 4.2)
	CT pELISA	75	89	31	709	29	693		45.7 (38.2 to 55.3)	4.1 (3.0 to 5.4)	3.0 (2.6 to 3.4)
Narvanen[268]	CT-EIA	69	13	150	2				84.2 (75.5 to 91.2)	1.3 (0.2 to 3.6)	6.3 (4.8 to 8.2)

FN, false negative; FP, false positive; TN, true negative; TP, true positive.
The analysis of the Morre study assumes that 'grey' samples were equally likely to be TPs or TNs. The analysis of the Wills *et al.*[269] study assumes that 18 paediatric samples that were either included (a) or excluded (b) from the calculation of FP rate were equally likely to be TPs or TNs.

Appendix 17 WinBUGS code applied to serology data

```
model {

# cycles through the different combinations of input parameters
for (a in 1:numcombs) {
 for (i in 1:17) {
  for (g in 1:2) {
   r[i,a,g] ~ dbin(p[g,a,i],n[i,a,g])
   rhat[i,a,g] <- p[g,a,i]*n[i,a,g]
   dev[i,a,g] <- 2 * (r[i,a,g] * (log(r[i,a,g])-log(rhat[i,a,g])) +
                     (n[i,a,g]-r[i,a,g]) * (log(n[i,a,g]-r[i,a,g]) -
                      log(n[i,a,g]-rhat[i,a,g])))
   }
  sumdev1[i,a] <- sum(dev[i,a,])
  p[1,a,i] <- pi[1,a] * se[i,a] + (1-pi[1,a]) * fp[i,a]
  p[2,a,i] <- pi[3,a]*sec[i,a] + pi[2,a] * se[i,a] + pi[4,a] *
fp[i,a]
  logit(se[i,a]) <- lse[i,a]
  logit(sec[i,a]) <- lse[i,a] + dse[a]
  lse[i,a] ~ dnorm(0,.01)
  lfp[i,a] <-  lse[i,a] - res[test[i],a]
  logit(fp[i,a]) <- lfp[i,a]
  }
 for (j in 1:5) {
  res[j,a] ~ dnorm(mean[a],prec[a])I(0,)
  }
 z[a] ~ dbeta(1,1)
 x[a] ~ dbeta(1,1)
 # Ever exposed in control group
 pi[1,a] <- ctprev[a]
 # Ever exposed in Non CT caused TFI
 pi[2,a] <- (1 - pi[3,a]) * (pi[1,a] + z[a] * (1 - pi[1,a]))
 # Proportion of TFI caused by C
 pi[3,a] <- x[a]
 # the negatices in the tfi group
 pi[4,a] <- (1 - pi[3,a]) * ((1 - z[a]) * (1 - pi[1,a]))
 # sum check
 pi[5,a] <- pi[2,a] + pi[3,a] + pi[4,a]
 sumdev2[a] <- sum(sumdev1[,a])
 }
```

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library